

APPLICANT: KHAN, Farrah A.
APPLICANT: YAO, Monique G.
APPLICANT: EMERLING, Brooke M.
TITLE OF INVENTION: INTRACELLULAR SIGNALING MOLECULES
FILE REFERENCE: PF-0868 USN
CURRENT APPLICATION NUMBER: US/10/451,207
CURRENT FILING DATE: 2003-06-18
PRIOR APPLICATION NUMBER: PCT/US01/50315
PRIOR FILING DATE: 2001-12-19
PRIOR APPLICATION NUMBER: US 60/257,804
PRIOR FILING DATE: 2000-12-21
PRIOR APPLICATION NUMBER: US 60/260,102
PRIOR FILING DATE: 2001-01-05
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PERL Program
SEQ ID NO 2
LENGTH: 670
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. US20040038267A1 5560161CD1
US-10-451-207-2

Query Match 91.4%; Score 3449.5; DB 12; Length 670;
Best Local Similarity 92.7%; Pred. No. 1.9e-293;
Matches 670; Conservative 0; Mismatches 0; Indels 53; Gaps 1;
QY 1 MILEERPDGAGAGESPRLOISRRKPKTRVSSLRGRREGRLDVCAMGRCVRHGEDPVR 60
DB 1 MILEERPDGAGAGESPRLO 20
QY 61 VHVGPMPQLHAGVCDLSLTQCCGLOSRRAQIHQOIKDELQMTGAEENLYRATSNRVR 120
DB 21 -----GCDLSLTQCCGLOSRRAQIHQOIKDELQMTGAEENLYRATSNRVR 67
QY 121 ETVALELSYVNSNQLKKEELEELSGGVDPRHSGSEAVTVPMIPLGLKETKELDWSTPLK 180
DB 68 ETVALELSYVNSNQLKKEELEELSGGVDPRHSGSEAVTVPMIPLGLKETKELDWSTPLK 127
QY 181 ELISVHFGEDEGASVEARELEALQAMRTSPRNSGLEGLELLTAYNQLCFDARFLTPAR 240
DB 128 ELISVHFGEDEGASVEARELEALQAMRTSPRNSGLEGLELLTAYNQLCFDARFLTPAR 187
QY 241 SLGLFFHWYDSLTGVPQAQORALAPEKGSVLFNICALHTQIGARDSCCTGARRAMEAFQ 300
DB 188 SLGLFFHWYDSLTGVPQAQORALAPEKGSVLFNICALHTQIGARDSCCTGARRAMEAFQ 247
QY 301 RAAGAFSLLENFSAHPSPDMSASLCALEQLMAQAQECVFEGLSPASMAPQDCCLAQL 360
DB 248 RAAGAFSLLENFSAHPSPDMSASLCALEQLMAQAQECVFEGLSPASMAPQDCCLAQL 307
QY 361 RLAGEAQAQAAEYRLVHRTWAOQPVHDYVPSWTALVHVVAEYFRSLAHVHVAALCDGS 420
DB 308 RLAGEAQAQAAEYRLVHRTWAOQPVHDYVPSWTALVHVVAEYFRSLAHVHVAALCDGS 367
QY 421 PATEGELPHEQVLOFPPTSSKPRGPVLPOLBEERRQLGKAHLKRALTLGQEEALRLHALC 480
DB 368 PATEGELPHEQVLOFPPTSSKPRGPVLPOLBEERRQLGKAHLKRALTLGQEEALRLHALC 427
QY 481 RVUREVDLRAVISQTLQORSLAKYAELEDREDDCEAAEAPDIQPKTHQKPEARMPLISQ 540
DB 428 RVUREVDLRAVISQTLQORSLAKYAELEDREDDCEAAEAPDIQPKTHQKPEARMPLISQ 487
QY 541 KGPDIHRLGFLSVFSKKNRRLVGPVHLTRGGGGLTLRGDSPVLIAAVIPGSAQAAA 600
DB 488 KGPDIHRLGFLSVFSKKNRRLVGPVHLTRGGGGLTLRGDSPVLIAAVIPGSAQAAA 547
QY 601 GLKEGDYIVSVNGOPCKWRHAEVWTELKAAAGAGASLOVSVLLPSRLPSLGDPRVLL 660
DB 548 GLKEGDYIVSVNGOPCKWRHAEVWTELKAAAGAGASLOVSVLLPSRLPSLGDPRVLL 607
QY 661 GPRGLLRSQREHGCKTFASTWASPRPLNWSRKAQQGKTGGCPQPCAPVKPAPSSSLKHP 720

DB 608 GPRGLLRSQREHGCKTFASTWASPRPLNWSRKAQQGKTGGCPQPCAPVKPAPSSSLKHP 667
QY 721 GWP 723
DB 568 GWP 670
RESULT 4
US-10-442-944-318
Sequence 318, Application US/10112944
Publication No. US20040048249A1
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Yang, Yonghong
APPLICANT: Weng, Gezhi
APPLICANT: Zhang, Jie
APPLICANT: Ren, Feiyan
APPLICANT: Xue, Aidong J.
APPLICANT: Wang, Jian-Rui
APPLICANT: Wehrman, Tom
APPLICANT: Ghosh, Malabika
APPLICANT: Wang, Dunrui
APPLICANT: Zhao, Qing A.
APPLICANT: Wang, Zhiwei
TITLE OF INVENTION: No. US20040048249A1el Nucleic Acids and
TITLE OF INVENTION: Secreted Polypeptides
FILE REFERENCE: 805A
CURRENT APPLICATION NUMBER: US/10/112,944
PRIOR FILING DATE: 2002-03-28
PRIOR APPLICATION NUMBER: US 09/488,725
PRIOR FILING DATE: 2000-01-21
PRIOR APPLICATION NUMBER: US 09/491,404
PRIOR FILING DATE: 2000-01-25
PRIOR APPLICATION NUMBER: US 09/496,914
PRIOR FILING DATE: 2000-02-03
PRIOR APPLICATION NUMBER: US 09/515,126
PRIOR FILING DATE: 2000-02-28
PRIOR APPLICATION NUMBER: US 09/519,705
PRIOR FILING DATE: 2000-03-07
PRIOR APPLICATION NUMBER: US 09/540,217
PRIOR FILING DATE: 2000-03-31
PRIOR APPLICATION NUMBER: US 09/552,929
PRIOR FILING DATE: 2000-04-18
PRIOR APPLICATION NUMBER: US 09/577,408
PRIOR FILING DATE: 2000-05-18
NUMBER OF SEQ ID NOS: 924
SOFTWARE: pt_FL_genes Version 5.0
SEQ ID NO 318
LENGTH: 669
TYPE: PRT
ORGANISM: Homo sapiens
US-10-112-944-318

Query Match 91.0%; Score 3435; DB 12; Length 669;
Best Local Similarity 92.5%; Pred. No. 3.6e-292;
Matches 669; Conservative 0; Mismatches 0; Indels 54; Gaps 2;
QY 1 MILEERPDGAGAGESPRLOISRRKPKTRVSSLRGRREGRLDVCAMGRCVRHGEDPVR 60
DB 1 MILEERPDGAGAGESPRLO 20
QY 61 VHVGPMPQLHAGVCDLSLTQCCGLOSRRAQIHQOIKDELQMTGAEENLYRATSNRVR 120
DB 21 -----GCDLSLTQCCGLOSRRAQIHQOIKDELQMTGAEENLYRATSNRVR 67
QY 121 ETVALELSYVNSNQLKKEELEELSGGVDPRHSGSEAVTVPMIPLGLKETKELDWSTPLK 180
DB 68 ETVALELSYVNSNQLKKEELEELSGGVDPRHSGSEAVTVPMIPLGLKETKELDWSTPLK 127
QY 181 ELISVHFGEDEGASVEARELEALQAMRTSPRNSGLEGLELLTAYNQLCFDARFLTPAR 240
DB 128 ELISVHFGEDEGASVEARELEALQAMRTSPRNSGLEGLELLTAYNQLCFDARFLTPAR 187

QY 241 SLGLFFHYDYSITGVPAQORALAFKSGVLEFNIGALHTQIGARODRSCCTEGARRAMEAFQ 300
 DB 188 SLGLFFHYDYSITGVPAQORALAFKSGVLEFNIGALHTQIGARODRSCCTEGARRAMEAFQ 247
 QY 301 RAAAGAFSLRENFHSHAPSPDMSAASLCALQOLMMAQAQECVFEGLSPDASMAPDCLAQ 360
 DB 248 RAAAGAFSLRENFHSHAPSPDMSAASLCALQOLMMAQAQECVFEGLSPDASMAPDCLAQ 307
 QY 361 RLAQEAQVAAEYRLVHRTMAQPPVHDVYVSWTALVHVKAIEYFRSLAHYHVAMALCDGS 420
 DB 308 RLAQEAQVAAEYRLVHRTMAQPPVHDVYVSWTALVHVKAIEYFRSLAHYHVAMALCDGS 367
 QY 421 PATEGELPHEQVFLQPPPTSSKPRGVLPOBLEERROLGKAHLKRAILGQEEALRLHALC 480
 DB 368 P-TEGELPHEQVFLQPPPTSSKPRGVLPOBLEERROLGKAHLKRAILGQEEALRLHALC 426
 QY 481 RVLREVDLLRAVISQTLQSLAKYAELEDDDFCEAAEAPDIQPKTHQKPEARMPLRSQ 540
 DB 427 RVLREVDLLRAVISQTLQSLAKYAELEDDDFCEAAEAPDIQPKTHQKPEARMPLRSQ 486
 QY 541 KGPDIFHRLGFLSVFSAKNRWLVGPVHLTGEGGFLTLRGDSPVLIAAVIPGSOQAAAA 600
 DB 487 KGPDIFHRLGFLSVFSAKNRWLVGPVHLTGEGGFLTLRGDSPVLIAAVIPGSOQAAAA 546
 QY 601 GLKEGDYIVSVNGQPCRWHRHAEVVTTELKAAAGEAGASLQVVSLLPSSRLSLGDRRPVLL 660
 DB 547 GLKEGDYIVSVNGQPCRWHRHAEVVTTELKAAAGEAGASLQVVSLLPSSRLSLGDRRPVLL 606
 QY 661 GPRGLLRSOREHGCKTPASTWASPRLLNWSRKAQOQKTCGCPQCAPVXPAPPSLKHP 720
 DB 607 GPRGLLRSOREHGCKTPASTWASPRLLNWSRKAQOQKTCGCPQCAPVXPAPPSLKHP 666
 QY 721 GWP 723
 DB 667 GWP 669

RESULT 5

US-10-697-266-4
 ; Sequence 4, Application US/10697266
 ; Publication No. US20040067522A1
 ; GENERAL INFORMATION:
 ; APPLICANT: RUSCH, Douglas et al
 ; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
 ; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
 ; FILE REFERENCE: CL001099-CIP-DIV2
 ; CURRENT APPLICATION NUMBER: US/10/697,266
 ; PRIOR FILING DATE: 2003-10-31
 ; PRIOR APPLICATION NUMBER: 10/274,878
 ; PRIOR FILING DATE: 2002-10-22
 ; PRIOR APPLICATION NUMBER: 09/849,634
 ; PRIOR FILING DATE: 2001-05-07
 ; PRIOR APPLICATION NUMBER: 09/773,371
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 4
 ; LENGTH: 634
 ; TYPE: PRT
 ; ORGANISM: Mus musculus
 US-10-697-266-4

Query Match 61.8%; Score 2331; DB 12; Length 634;
 Best Local Similarity 67.0%; Pred. No. 2.7e-195;
 Matches 473; Conservative 49; Mismatches 112; Indels 72; Gaps 6;
 QY 1 MILEERPDCAGAGESPRLQISRRKPRKTRVSSLRGRREGRLDVCWARGCRVHRGEDPVR 60
 DB 1 MILEERPDCAGAGESPRLQISRRKPRKTRVSSLRGRREGRLDVCWARGCRVHRGEDPVR 26
 QY 61 VHVGPMPQLHVGCDLSITQCGQLQSRRAQIHQIDKELQMTGAENLYRATSNRVR 120

DB 27 -----KYGSGFVNQPCQLQSHRARLHQIISKELRMTGAENLYRATSNRVR 74
 QY 121 ETVALELSYNSNLOLLKEELEISGSDVDPGRHGESEAVTVMIPGLKTKELDWSTPLK 180
 DB 75 ETVALELSYNSNLOLLKEELEISGSDVDPGRHGESEAVTVMIPGLKTKELDWSTPLK 134
 QY 181 ELISVHGEDGASVEAREIRELEALROAMRTSPSRNESGLELLTAYYNOLCFDARFLPAR 240
 DB 135 ELISEHGEDGTSEFEIQELEDRLQATRTSPSRNESGLELLTAYYNOLCFDARFLPAR 194
 QY 241 SLGLFFHYDYSITGVPAQORALAFKSGVLEFNIGALHTQIGARODRSCCTEGARRAMEAFQ 300
 DB 195 SPGLLFFHYDYSITGVPAQORALAFKSGVLEFNIGALHTQIGARODRSCCTEGARRAMEAFQ 254
 QY 301 RAAAGAFSLRENFHSHAPSPDMSAASLCALQOLMMAQAQECVFEGLSPDASMAPDCLAQ 360
 DB 255 RAAAGAFSLRENFHSHAPSPDMSAASLCALQOLMMAQAQECVFEGLSPDASMAPDCLAQ 314
 QY 361 RLAQEAQVAAEYRLVHRTMAQPPVHDVYVSWTALVHVKAIEYFRSLAHYHVAMALCDGS 420
 DB 315 RLAQEAQVAAEYRLVHRTMAQPPVHDVYVSWTALVHVKAIEYFRSLAHYHVAMALCDGS 374
 QY 421 PATEGELPHEQVFLQPPPTSSKPRGVLPOBLEERROLGKAHLKRAILGQEEALRLHALC 480
 DB 375 PA-KGELARQEHV-FQSTPHEPLGPTLPQHPEDRKLAKAHLKRAILGQEEALRLHALC 432
 QY 481 RVLREVDLLRAVISQTLQSLAKYAELEDDDFCEAAEAPDIQPKTHQKPEARMPLRSQ 540
 DB 433 RVLREVDLLRAVISQTLQSLAKYAELEDDDFCEAAEAPDIQPKTHQKPEARMPLRSQ 483
 QY 541 KGPDIFHRLGFLSVFSAKNRWLVGPVHLTGEGGFLTLRGDSPVLIAAVIPGSOQAAAA 600
 DB 484 -----GPLSVFSTKNRWLVGPVHMTGEGGFLTLRGDSPVLIAAVIPGSOQAAAA 534
 QY 601 GLKEGDYIVSVNGQPCRWHRHAEVVTTELKAAAGEAGASLQVVSLLPSSRLSLGDRRPVLL 660
 DB 535 GLKEGDYIVSVNGQPCRWHRHAEVVTTELKAAAGEAGASLQVVSLLPSSRLSLGDRRPVLL 594
 QY 661 GPRGLLRSOREHGCKTPASTWASPRLLNWSRKAQOQKTCGCPQCAPVXPAPPSLKHP 706
 DB 595 W-----NORECGFETPMPTRTRPWPILGWSRKNKQKGTGSHDPDC 634

RESULT 6

US-10-274-878-4
 ; Sequence 4, Application US/10274878
 ; Publication No. US20030049792A1
 ; GENERAL INFORMATION:
 ; APPLICANT: RUSCH, Douglas et al
 ; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
 ; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
 ; FILE REFERENCE: CL001099-CIP-DIV
 ; CURRENT APPLICATION NUMBER: US/10/274,878
 ; PRIOR FILING DATE: 2002-10-22
 ; PRIOR APPLICATION NUMBER: 09/849,334
 ; PRIOR FILING DATE: 2001-05-07
 ; PRIOR APPLICATION NUMBER: 09/773,371
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 4
 ; LENGTH: 634
 ; TYPE: PRT
 ; ORGANISM: Mus musculus
 US-10-274-878-4

Query Match 61.8%; Score 2331; DB 14; Length 634;
 Best Local Similarity 67.0%; Pred. No. 2.7e-195;
 Matches 473; Conservative 49; Mismatches 112; Indels 72; Gaps 6;
 QY 1 MILEERPDCAGAGESPRLQISRRKPRKTRVSSLRGRREGRLDVCWARGCRVHRGEDPVR 60

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 13, 2004, 11:14:21 ; Search time 83 Seconds
(without alignments)
2748.430 Million cell updates/sec

Title: US-10-697-266-2
Perfect score: 3774
Sequence: 1 MILEERPDGAGAESPRLO.....QPCAPVKPAPPSLKHGWP 723

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL 25:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3449.5	91.4	670	Q8TAV1	Q8tav1 homo sapien
2	1847	48.9	410	Q96PV9	Q96pv9 homo sapien
3	1496	39.6	451	Q80WU2	Q80wu2 mus musculus
4	1316	34.9	683	Q803B1	Q803b1 brachydanio
5	1284	34.0	686	Q8BXG3	Q8bxg3 canis famil
6	1255	33.3	686	Q8NE33	Q8ne33 homo sapien
7	1255	33.3	686	Q8IUC4	Q8iuc4 homo sapien
8	1242	32.9	686	Q9DEN2	Q9den2 mus musculus
9	1237	32.8	657	Q8N3T7	Q8n3t7 homo sapien
10	1228	32.5	686	Q8BWR8	Q8bwr8 mus musculus
11	1185.5	31.4	685	Q96RU1	Q96ru1 homo sapien
12	1023	27.1	718	Q9XY9	Q9xy9 drosophila
13	887.5	23.5	687	Q9VXP2	Q9vxp2 drosophila
14	394	10.4	298	Q8N9D6	Q8n9d6 homo sapien
15	387	10.3	1421	Q9P257	Q9p257 homo sapien
16	380.5	10.1	1636	Q9H3S7	Q9h3s7 homo sapien

17	335.5	8.9	794	5	Q8T7K0	Q8t7k0 dictyosteli
18	333	8.8	846	10	Q8HLH8	Q8hlh8 arabidopsis
19	295.5	7.8	883	10	Q8LNU2	Q8lnu2 oryza sativ
20	292.5	7.8	773	3	Q8WZL4	Q8wzl4 yarrowia li
21	289	7.7	816	10	Q9X156	Q9xi56 arabidopsis
22	285.5	7.6	775	3	Q13783	Q13783 schizosacch
23	284.5	7.5	846	5	Q8I114	Q8i114 caenorhabdi
24	284.5	7.5	882	5	Q8I115	Q8i115 caenorhabdi
25	284	7.5	465	13	Q7SY03	Q7sy03 brachydanio
26	276.5	7.3	498	5	Q7YXC7	Q7yxc7 caenorhabdi
27	266.5	7.1	872	11	Q8Y09	Q8y09 mus musculu
28	262.5	7.0	1833	5	Q9VUH6	Q9vuh6 drosophila
29	262.5	7.0	1838	5	Q960G3	Q960g3 drosophila
30	255.5	6.8	847	3	P79020	P79020 emericeila
31	239	6.3	836	5	Q9VB05	Q9vb05 drosophila
32	232	6.1	661	3	Q12033	Q12033 saccharomyc
33	224.5	5.9	266	5	Q9NA90	Q9na90 caenorhabdi
34	203.5	5.4	785	3	Q9VU12	Q9vu12 candida alb
35	165.5	4.4	1802	5	Q9VQU8	Q9vqu8 drosophila
36	165	4.4	1190	5	Q9V562	Q9v562 drosophila
37	165	4.4	1205	5	Q8MKS8	Q8mks8 drosophila
38	162	4.3	1386	5	Q8MKS7	Q8mks7 drosophila
39	162	4.3	1407	5	Q8IMD0	Q8imd0 drosophila
40	155.5	4.1	942	4	Q8IUV5	Q8iuu5 homo sapien
41	155.5	4.1	3753	2	Q846W6	Q846w6 streptomyce
42	151	4.0	1024	11	Q9ESJ5	Q9esj5 mus musculu
43	149	3.9	799	13	Q7SYH9	Q7syh9 gallus gall
44	147	3.9	1151	2	Q84EM0	Q84em0 rhodobacter
45	144	3.8	966	4	Q8NDG1	Q8ndg1 homo sapien

ALIGNMENTS

RESULT 1

ID	Q8TAV1	PRELIMINARY;	PRT;	670 AA.
AC	Q8TAV1			
DT	01-JUN-2002 (Tremblrel. 21, Created)			
DT	01-JUN-2002 (Tremblrel. 21, Last sequence update)			
DT	01-JUN-2003 (Tremblrel. 24, Last annotation update)			
DE	Similar to GTP-rho binding protein 1.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Blood;			
RA	Strausberg R.;			
RL	Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.			
CC	-!- SIMILARITY: CONTAINS 1 PDZ/DHR DOMAIN.			
DR	EMBL; BC025767; AAH25767.1; -			
DR	GO; GO:0005622; C:intracellular; IEA.			
DR	GO; GO:0007242; P:intracellular signaling cascade; IEA.			
DR	InterPro; IPR004328; BR01.			
DR	InterPro; IPR001478; PDZ.			
DR	InterPro; IPR000861; REM_repeat.			
DR	Pfam; PF03097; BR01; 1.			
DR	Pfam; PF02185; HR1; 1.			
DR	Pfam; PF00595; PDZ; 1.			
DR	SMART; SM00074; HR1; 1.			
DR	SMART; SM00228; PDZ; 1.			
DR	PROSITE; PS50106; PDZ; 1.			
SQ	SEQUENCE 670 AA; 73589 MW; CE8C01B716465FF0 CRC64;			

Query Match 91.4%; Score 3449.5; DB 4; Length 670;
Best Local Similarity 92.7%; Pred. No. 1.2e-239;
Matches 670; Conservative 0; Mismatches 0; Indels 53; Gaps 1;

Oy 1 MILEERPDGAGAESPRLOISRRKPKTRVSSLRGREGLRDVCANRGCKVHGDFVR 60
Db 1 MILEERPDGAGAESPRLO-----

DR GO: 0005622; C:intracellular; IEA.
 DR GO: 0007165; P:signal transduction; IEA.
 DR InterPro: IPR004328; BRO1.
 DR InterPro: IPR000861; REM_repeat.
 DR Pfam: PF03097; BRO1; 1.
 DR Pfam: PF02185; HRI; 1.
 DR SMART: SM00074; HRI; 1.
 KW Hypothetical protein.
 FT NON TER 1
 SQ SEQUENCE 410 AA; 73920DC672235189 CRC64;
 Query Match 48.9%; Score 1847; DB 4; Length 410;
 Best Local Similarity 81.4%; Pred. No. 1.3e-124;
 Matches 377; Conservative 2; Mismatches 22; Indels 62; Gaps 4;
 QY 4 EERPDGAGAGEEPRLOISRRKPRKTRVSSILRGRREGRLDYCAWRGCRVHRGDEPVRVHV 63
 Db 1 EERPDGAGAGEEPRLO-----17
 QY 64 GPMNPQLHAVGCDLSLTQICGQLOSRAQHQIQDIDKELQMTGAENLYRATSNRRVETV 123
 Db 18 -----GCDLSLTQICGQLOSRAQHQIQDIDKELQMTGAENLYRATSNRRVETV 67
 QY 124 ALELSYNSNLIQLKEELEELSGGVDPGRHSGSEAVTPMPLGLKETKELDWSPLKELI 183
 Db 68 ALELSYNSNLIQLKEELEELSGGVDPGRHSGSEAVTPMPLGLKETKELDWSPLKELI 127
 QY 184 SVHFGEDGASYEAEIRLEALRQAMRTSPSRNESGLELLTAYYNQICFLDARFLTPARSLG 243
 Db 128 SVHFGEDGASYEAEIRLEALRQAMRTSPSRNESGLELLTAYYNQICFLDARFLTPARSLG 187
 QY 244 LFFHWYDLSLTGVPAQORALAFKESVLEPNICGALHTQIGARODRSCTEGARRAMEAFQRAA 303
 Db 188 LFFHWYDLSLTGVPAQORALAFKESVLEPNICGALHTQIGARODRSCTEGARRAMEAFQRAA 247
 QY 304 GAFSLIRENFHSPDMSAASLCALEQLMMAQAQECVFEGISPPASMAPQDCLAQLRIA 363
 Db 248 GAFSLIRENFHSPDMSAASLCALEQLMMAQAQECVFEGISPPASMAPQDCLAQLRIA 307
 QY 364 QEAQAQVAAEYRLVHRTMAQPPVHDYVPVSWTALVHVKAIEYFRSLAHYHVAMALCDGSPAT 423
 Db 308 QEAQAQVAAEYRLVHRTMAQPPVHDYVPVSWTALVHVKAIEYFRSLAHYHVAMALCDGSPAT 466
 QY 424 EGELTHQVFLQPTSSKPRGPVLPOELEERRQLGKAHLKEA 466
 Db 365 -RECPHLEPMVLRP-----PRAG-LTASLSPSSDRGRAPHARA 401

RESULT 3

Q80WU2 PRELIMINARY; PRT; 451 AA.
 AC Q80WU2;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Similar to rhophilin, Rho GTPase binding protein 1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Brain;
 RX MEDLINE=22398257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Bhat N.K.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hong L.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Frange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RESULT 2

Q96PV9 PRELIMINARY; PRT; 410 AA.
 AC Q96PV9;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Hypothetical protein KIAA1929 (Fragment).
 GN KIAA1929.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=21456161; PubMed=11572484;
 RA Nagase T., Kikuno R., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. XXI.
 RT The complete sequences of 60 new cDNA clones from brain which code for
 RT large proteins.";
 RL DNA Res. 8:179-187(2001).
 DR EMBL; AB067516; BAB67822.1; -.

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Ketterman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences,"
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RA Strausberg R.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC052010; AAH52010.1; -;
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR004328; BRO1.
DR InterPro; IPR000861; REM_repeat.
DR Pfam; PF03097; BRO1; 1.
DR Pfam; PF02185; HRI; 1.
DR SMART; SM00074; HRI; 1.
DR PROSITE; PS00074; HRI; 1.
SQ SEQUENCE 451 AA; 49072 MW; D69B6A232CA19018 CRC64;

Query Match 39.6%; Score 1496; DB 11; Length 451;
Best Local Similarity 67.3%; Pred. No. 2.7e-99;
Matches 309; Conservative 28; Mismatches 74; Indels 48; Gaps 4;

QY 1 MILERRPDGAGGESPRIOISRRKPRKTRVSSLRGRREGRLDVCAMWRCRVHRGDPVR 60
DB 1 MILERRPDGQGGERSRRPQDD-----GSIR----- 26

QY 61 VHVGPMPQLHVGCDLSLTQICGQLOSRRAHQHIDKELQWRTGAENLYRATSNRYR 120
DB 27 -----KGYGSFVQNPQPGQLQSHRARLHQHISKELRMRTGAENLYRATSNRYR 74

QY 121 ETVALELSVNSNLQKLEELSGVDPGRHGSFAVTPMILGLKETKELDWSLTK 180
DB 75 ETVALELSVNSNLQKLEELSLTSVDVDQPEGETITPILGLKETKELDWSLTK 134

QY 181 ELISVHFGEDGASYEARELEALRQAMRTPSRNFSGLLELTAYYNQLCFDLARLTAR 240
DB 135 ELISVHFGEDGTSFETIQELEDLAQATPTPSRDEAGDLLAAYISQLCFDLARFSPSR 194

QY 241 SGLFPEHWDLSLTGVPQAORALAFKGSVLFNIGALHTQIGARQDRCTEGARRAMEAFQ 300
DB 195 SPGLLFHWDLSLTGVPQAORALAFKGSVLFNIGALHTQIGARQDRCTEGTNHAAEFQ 254

QY 301 RAAGAFSLLENFNSHAPSPDMSAASLCALQALMAQACQCEGSLSPASMAPQCLAQ 360
DB 255 RAAGAFSLLENFNSHAPSPDMSAASLMLEQUNTAQACQCEFKGULLPASATPDICPDQL 314

QY 361 RLAEAAQAAVAFYRLVHRTMAQPPVHDYVPVSWTALVHVKAIFYRSLAHYHVMALCDGS 420
DB 315 QLAQAAQAAVAFYGLVHRAMAQPPVHDYVPVSWTALVHVKAIFYRSLAHYHVMALCESH 374

QY 421 PATEGELTHQVFLQPTTSKPRGPVLPOELEERRQLG 459
DB 375 PA-KGELARQEHVF-QPSTPHEPLGFTLPOHPEDRRKLG 411

RESULT 4

Q803B1
ID Q803B1 PRELIMINARY; PRT; 683 AA.
AC Q803B1;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similar to rhophilin-2.
OS Brachydanio rerio (zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC044556; AAH44556.1; -;
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR InterPro; IPR004328; BRO1.
DR InterPro; IPR001478; PDZ.
DR InterPro; IPR000861; REM_repeat.
DR Pfam; PF03097; BRO1; 1.
DR Pfam; PF02185; HRI; 1.
DR Pfam; PF00595; PDZ; 1.
DR SMART; SM00074; HRI; 1.
DR SMART; SM00228; PDZ; 1.
DR PROSITE; PS0106; PDZ; 1.
SQ SEQUENCE 683 AA; 77001 MW; B06F4D2B088FCB52 CRC64;

Query Match 34.9%; Score 1316; DB 13; Length 683;
Best Local Similarity 44.3%; Pred. No. 4.5e-86;
Matches 271; Conservative 122; Mismatches 205; Indels 14; Gaps 4;

QY 41 LRDVCAWRCRVHRGDPVRVHVGPMPQLHVGCDLSLTQICGQLOSRRAHQHIDKE 100
DB 1 MTDLTLPNGCKSN-----GPVNGYFKGCNPNFAQTGSKLONKRAVLNKLIIQ 50

QY 101 LQWRTGAENLYRATSNRYRRTVALELSVNSNLQKLEELSGVDPGRHGSFAVTP 160
DB 51 MMRAGAENLLKATSNKRVQVLELSYVNSDLQLLMEQLGLNSSVEYVQVQSSSI 110

QY 161 PMIPGLKETKELDWSLTKELISVHFGEDGASYEARELEALRQAMRTPSRNFSGL 220
DB 111 PLIPGLKETKDVDPVPLKDFILEHSEDGNSFNQYIDDLMDLQACQTPSRNNSGVDL 170

QY 221 LPAYYNQLCFDLARLTARLTARSLGLFHHWDLSLTGVPQAORALAFKGSVLFNIGALHTOI 280
DB 171 IANYPSQLSFLTRFPSTPQIGIFFTWVDSFTGMPVCNNISLEKASMLFNMAALYSQI 230

QY 281 GARQDRCTEGARRAMEAFQRAAGAFSLLENFNSHAPSPDMSAASLCALQALMAQAC 340
DB 231 GTRADRQTLAGLEDAIAAFQKSAAGVHLHLLKETFTHTPTPSYDMSAPMLSLFRMLAQAC 290

QY 341 VPEGLSPASMAPQCLAQALRAQAAQVAAEYRLVHRTMAQPPVHDYVPVSWTALVHV 400
DB 291 LFEQITLPGIRNEFFCL--LWQAQAAKVAETYAQVHOSMIETPTIKNVFFWTMSQLK 348

QY 401 AEYFSLAHYHVMALCDGSPATEGELTHQVFLQ-PFTSSKPRGPV-LPOELEERRQL 458
DB 349 INHYNLSAHYFVSTALLDHLQNPSSDEKQKALSQLYDAMPEGSPDLILKNKDERRI 408

QY 459 GKALHRAILQGEALRLHALCRVLREVDLRAVLSQTLQSLAKYAELEDREDDFCEAAE 518
DB 409 GKALHRSIMGHEAITHCRRLHQLDILSLIRASLNRSLSLTKFEQNDKEDFTDYL 468

QY 519 APDIQPKTHQKPEARMPLRSQKGPDIPIHRLGPLSVFSAKNRWRLVGPVHLTRGEGGCL 578
DB 469 APDIISKTEKAEAEIIPAATKVKTDLFORLGPLSVFSAKQRTAPRTIRLILQDRDLGF 528

QY 579 TLRGDSPVLIAAVTPGSOAAAGLKEGDYIVSVNGQPCRRWRHAEVVTELKAAAGAGASL 638
DB 529 TLKGDAPVQIQSLDPLCPAAAGLKEGDYLVAVGDTCKWNGVSDVMKLLKDVDERGINI 588

QY 639 QVWSLLPSSRLP 650
DB 589 RVVSMMDSSSQP 600

RESULT 5

Q8HXG3
ID Q8HXG3 PRELIMINARY; PRT; 686 AA.


```

QY 194 YEAFIRELEALROAMRTSPRESNGLLELTAYYNOLCFDARFLTPARSLGLFFHWYDSLT 253
D 194 YEDDIADLMQDRACRTSPDRDEAGVELLMSTFIQLGFEVSFFPPTHMGLLFTWYDSFT 203
QY 254 GVPAAQRALAPEKGSVLFNIGALHTQIGARQDRSCTEGARRAMEAFQRAAGAFSLLENF 313
D 204 GVPVSQQTLLLEKASVLFNIGALYTIQIGTRCNROTQAGLESADVDAFQRAAGVNLKETF 263
QY 314 SHAPSPMSAASLALCALBOLMAQAQECVFEGSLSPASMAPQCLALQALAEAAQVAAY 373
D 264 THTESYDMSPAVLSLVNMLAQAESVFVKCLPG--IQNEFFVLVKAQEAQVAAY 321
QY 374 RLVRHTMAQPVHDVVPVSWTALHVHKAIFYERSLAHYHVMALCDG--SPATEGELPTHE 431
D 322 RQLHAAMSQEVKENIPYSWASVAYVKAHYGALAHYFAATLLIDHQLKPGADE--HQ 378
QY 432 QVFLQPPTSSKPRGVLVPLE--ERRQLGKHLKRAILGOEALRLHALCRVLRVDL 488
D 379 EKCLSQLYDRMEGWTPLATLKNAGQVLLGKHLHRAIGFHEESLREANLCKKLRIQV 438
QY 489 LRAVISQTLQSLAKYABLDREDDPCEAAABDIOPTKQKPEAMPRLSQKGPDIHR 548
D 439 LEDVLSAAHQRTQLKHTQHRREDLNLIDAPVLPKTEREVKITPFDPSKVTVTFQK 498
QY 549 LGPLSVFSKRNRLVGPVHLTRGEGFGLTLRGDSPVLIAAVIPGSAAGLKEGDIYI 608
D 499 LGPLSVFSKRNWSPRGHIFVTEGDLGFTLRGNTPVQVHFLDHCASLAGAKEDIYI 558
QY 609 VSVNGQPCRWHRHARVVTTELKAAGSAGASLQVSVLLPSS 647
D 559 VSIQGVCKWLTVSEVMKILKSGFGEVEMKVSVLLDST 597

RESULT 9
Q8N3T7 PRELIMINARY; PRT; 657 AA.
AC Q8N3T7;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Hypothetical protein (Fragment).
GN DKFZP761L1918.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ON NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Amygdala;
RA Ottenwaelder B., Obermaier B., Mewes H.W., Weil B., Wiemann S.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL831950; CAD38597.1; -.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR InterPro; IPR004328; BRO1.
DR InterPro; IPR001478; PDZ.
DR InterPro; IPR000861; REM_repeat.
DR Pfam; PF03097; BRO1; 1.
DR Pfam; PF02185; HRI; 1.
DR SMART; SM00074; HRI; 1.
DR SMART; SM00228; PDZ; 1.
KW Hypothetical protein.
FT NON_TER
SQ SEQUENCE 657 AA; 73850 MW; 562DEF16F238B39D CRC64;

Query Match 32.8%; Score 1237; DB 4; Length 657;
Best Local Similarity 46.1%; Pred. No. 2.1e-80;
Matches 262; Conservative 106; Mismatches 190; Indels 10; Gaps 5;

QY 85 QLOSRAQIHQQIDKELQMTGARNLYRATSNRNRVETVALELSYVNSNLQLLKEELEEL 144
D 6 KLQNGRAALNQQILKAVRMRTGAENLLKVATNSKVRQVRLLELSFVNSDLQMLKEELEGL 65

```

```

QY 145 SGGVDPQRHSGEAVTVPMIPGLKTKELDWSTPLKELISVHFGECDGASYEABIRELEAL 204
D 66 NISVGVYNTTEAFTIPLIPLGLKETKDQVDFAVVYDKDFILEHYSEGDYLYEIDEADLMDL 125
QY 205 ROAMRTSPRESNGLLELTAYYNOLCFDARFLTPARSLGLFFHWYDSLTGVPAAQRALAF 264
D 126 RQAKRTSPDRDEAGVELLMSTFIQLGFEVSFFPPTHMGLLFTWYDSLTGVPVQSNLL 185
QY 265 EKGSVLFNIGALHTQIGARQDRSCTEGARRAMEAFQRAAGAFSLLENFSHAPSPDMSAA 324
D 186 EKASVLFNTGALYTIQIGTRCDROTQAGLESADAFQRAAGVNLKDTFTHTPSYDMSPA 245
QY 325 SLCALBOLMAQAQECVFEGSLSPASMAPQCLALQALAEAAQVAAYRVLVHRTMAQPP 384
D 246 MLSVLVNMMLAQAESVFVKISLPG--IRNEFFMLVKVQAQEAQVGEYVQQLHAAMSQAP 303
QY 385 VHDVVPVSWTALHVHKAIFYERSLAHYHVMALCDG--SPATEGELPTHEQVFLQPPTSSK 442
D 304 VKENIPYSWASLACVKAHYAALAHYFAATLLIDHQLKPGTDLD--HQECLSLQLYDHM 360
QY 443 PRG--PVLP--QELERRQLGKHLKRAILGOEALRLHALCRVLRVDLRLRAVISQTLQ 499
D 361 PEGLTPLATLKNQDQRRQLGKSHLRRAMAHHEESVREASLCKLRSIEVLQKVLCAQER 420
QY 500 SLAKYAEALDRDDPCEAAABDIOPTKQKPEAMPRLSQKGPDIHRLGLPLSVFSAKN 559
D 421 SRLTYAQHEEDLNLIDAPVLPKTEREVKITPFDPSKVTVTFQKLGPLSVFSANK 480
QY 560 RWLGVPHLTRGEGFGLTLRGDSPVLIAAVIPGSAAGLKEGDIYVSVNGQPCRW 619
D 481 RWTTPRSRFTAEEDGDLGFTLRGNAPVQVHFLDPYCSASVAGAREGDIYVSVLQVDCRW 540
QY 620 RHAEVVTELKAAGSAGASLQVSVLLPSS 647
D 541 TLSEVMKILKSGFGEIEMKVSVLLDST 568

RESULT 10
Q8BWR8 PRELIMINARY; PRT; 686 AA.
AC Q8BWR8;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Similar to rhophilin-like protein.
GN RHPN2 OR D7ERTD784E.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
ON NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Liver;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
DR EMBL; AK050214; BAC34127.1; -.
DR MGD; MGI:1289234; Rhp2.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR InterPro; IPR004328; BRO1.
DR InterPro; IPR001478; PDZ.
DR InterPro; IPR000861; REM_repeat.
DR Pfam; PF03097; BRO1; 1.
DR Pfam; PF02185; HRI; 1.
DR Pfam; PF00595; PDZ; 1.
DR SMART; SM00074; HRI; 1.
DR SMART; SM00228; PDZ; 1.
DR PROSITE; PS50106; PDZ; 1.
SQ SEQUENCE 686 AA; 76905 MW; 49B67AA361FEDC58 CRC64;

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 13, 2004, 12:02:51 ; Search time 30 Seconds
(without alignments)
1244.186 Million cell updates/sec

Title: US-10-697-266-2
Perfect score: 3774
Sequence: 1 MILBERPDGAGGESPR1Q.....QPCAPVKPAPSSLKHPGWP 723

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3774	100.0	723	4	US-09-849-334-2
2	3774	100.0	723	4	US-10-274-878-2
3	2331	61.8	634	4	US-09-849-334-4
4	2331	61.8	634	4	US-10-274-878-4
5	268	7.1	921	4	US-09-800-729-199
6	264.5	7.0	868	4	US-09-800-729-106
7	155.5	4.1	942	4	US-08-685-852-3
8	134.5	3.6	450	2	US-08-665-037-2
9	134.5	3.6	450	2	US-08-666-067-2
10	134.5	3.6	450	2	US-08-732-870-2
11	134	3.6	1266	4	US-09-252-991A-30851
12	131.5	3.5	2293	3	US-09-368-590-2
13	129	3.4	751	4	US-09-252-991A-27424
14	128.5	3.4	658	4	US-09-252-991A-18483
15	128	3.4	864	4	US-09-883-096-2
16	125.5	3.3	635	4	US-09-252-991A-18485
17	123.5	3.3	1792	4	US-09-561-818A-4
18	123.5	3.3	1800	4	US-09-561-818A-8
19	123.5	3.3	1816	4	US-09-561-818A-2
20	123.5	3.3	1824	4	US-09-561-818A-6
21	123	3.3	609	4	US-09-252-991A-24893
22	123	3.3	643	4	US-09-252-991A-22490
23	122	3.2	358	4	US-09-740-027-4
24	121.5	3.2	729	4	US-09-252-991A-22350
25	120	3.2	1665	4	US-09-858-664A-2
26	120	3.2	1665	4	US-10-274-978-2
27	120	3.2	6095	3	US-09-144-085-2

28	119.5	3.2	223	4	US-09-124-238A-34	Sequence 34, Appl
29	119.5	3.2	223	4	US-09-721-975-34	Sequence 34, Appl
30	119.5	3.2	223	4	US-09-986-621-34	Sequence 34, Appl
31	119.5	3.2	457	4	US-09-124-238A-10	Sequence 10, Appl
32	119.5	3.2	457	4	US-09-721-975-10	Sequence 10, Appl
33	119.5	3.2	457	4	US-09-986-621-10	Sequence 10, Appl
34	119.5	3.2	819	4	US-09-651-656-15	Sequence 15, Appl
35	119.5	3.2	819	4	US-09-650-855-15	Sequence 15, Appl
36	118	3.1	1445	1	US-08-015-986A-2	Sequence 2, Appl
37	118	3.1	1445	2	US-08-446-363-2	Sequence 2, Appl
38	117.5	3.1	343	4	US-09-252-991A-24630	Sequence 24630, A
39	117.5	3.1	1050	3	US-09-045-632-50	Sequence 50, Appl
40	117.5	3.1	2504	4	US-08-851-567B-12	Sequence 12, Appl
41	117.5	3.1	2504	4	US-09-817-514A-8	Sequence 8, Appl
42	117	3.1	629	4	US-09-252-991A-31243	Sequence 31243, A
43	116.5	3.1	1739	4	US-09-976-594-76	Sequence 76, Appl
44	116	3.1	457	4	US-09-124-238A-1	Sequence 1, Appl
45	116	3.1	457	4	US-09-721-975-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-09-849-334-2
; Sequence 2, Application US/09849334
; Patent No. 6500655
; GENERAL INFORMATION:
; APPLICANT: RUSCH, Douglas et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001099-CIP
; CURRENT APPLICATION NUMBER: US/09/849,334
; CURRENT FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 723
; TYPE: PRT
; ORGANISM: Human
US-09-849-334-2

Query Match	100.0%	Score 3774;	DB 4;	Length 723;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 723;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
QY	1	MILBERPDGAGGESPR1QISRRKPKRTRVSSLRGREGLDVCAWRCGRVHRGEDPVR	60	
Db	1	MILBERPDGAGGESPR1QISRRKPKRTRVSSLRGREGLDVCAWRCGRVHRGEDPVR	60	
QY	61	VHVGPMPNQLHAVGCDLSLTQCCQLOSRRAQIHQCIDKELQMTGAENLYRATSNRVR	120	
Db	61	VHVGPMPNQLHAVGCDLSLTQCCQLOSRRAQIHQCIDKELQMTGAENLYRATSNRVR	120	
QY	121	ETVALEISYNSNQLKKELEELSGGVDPCRHSSEAVTPMIPGLKXETKELDMSTELK	180	
Db	121	ETVALEISYNSNQLKKELEELSGGVDPCRHSSEAVTPMIPGLKXETKELDMSTELK	180	
QY	181	ELISVHFGEDGASVEARELEALRQAMRTSPRNESGLELLTAYNQICFLDARFLTPAR	240	
Db	181	ELISVHFGEDGASVEARELEALRQAMRTSPRNESGLELLTAYNQICFLDARFLTPAR	240	
QY	241	SLGLFFHWYDSLTGVPQAQALAFKESVLFNICALHTQIGARQDRSCTEGARRAMEAFQ	300	
Db	241	SLGLFFHWYDSLTGVPQAQALAFKESVLFNICALHTQIGARQDRSCTEGARRAMEAFQ	300	
QY	301	RAAGAFSLLRENFHSPDMSAASLCALEQMAAQECVFEGSLSPASMAPQDCLAQL	360	
Db	301	RAAGAFSLLRENFHSPDMSAASLCALEQMAAQECVFEGSLSPASMAPQDCLAQL	360	
QY	361	PLAQEAQVAAYEELVHRTMAOPPHYVPSVTALVHVKAIEYFRSLAHYVAMALCDGS	420	
Db	361	PLAQEAQVAAYEELVHRTMAOPPHYVPSVTALVHVKAIEYFRSLAHYVAMALCDGS	420	

Db 361 RLAEAAQVAAEYRLVHRTWAQPPVHDYVPVSWTALVHVKAAYFRSLAHYHVAMALCDGS 420
QY 421 PATEGELPTEHQVFLQPTTSKPRGPVLPOELERRQLGKHLKRAILGQEEALRLHALC 480
Db 421 PATEGELPTEHQVFLQPTTSKPRGPVLPOELERRQLGKHLKRAILGQEEALRLHALC 480
QY 481 RVLRVDLLRAVISQTLQSLAKYAELEDRDDFCEAAEAPDIQPKTHQKPEARNPRLSQG 540
Db 481 RVLRVDLLRAVISQTLQSLAKYAELEDRDDFCEAAEAPDIQPKTHQKPEARNPRLSQG 540
QY 541 KGPDIHRLGFLSVFSAKNRWLVGPHVLTGRGGFGLTLRGDSPVLIAAVIPGSQAAAA 600
Db 541 KGPDIHRLGFLSVFSAKNRWLVGPHVLTGRGGFGLTLRGDSPVLIAAVIPGSQAAAA 600
QY 601 GLKEGDYIVSVNGQPCRWWRHAEVVTTELKAAGEAGASLOVVSLLPSSRLSLGDRRPVLL 660
Db 601 GLKEGDYIVSVNGQPCRWWRHAEVVTTELKAAGEAGASLOVVSLLPSSRLSLGDRRPVLL 660
QY 661 GPRGLLRQREHGCKTTPASTWASPRLLNWSRKAQCGKTCGCPQPCAPVKPAPPSLKH 720
Db 661 GPRGLLRQREHGCKTTPASTWASPRLLNWSRKAQCGKTCGCPQPCAPVKPAPPSLKH 720
QY 721 GWP 723
Db 721 GWP 723

RESULT 2
US-10-274-878-2
; Sequence 2, Application US/10274878
; Patent No. 6670163
; GENERAL INFORMATION:
; APPLICANT: RUSCH, Douglas et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001099-CIP-DIV
; CURRENT APPLICATION NUMBER: US/10/274,878
; PRIOR FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: 09/849,334
; PRIOR FILING DATE: 2001-05-07
; PRIOR APPLICATION NUMBER: 09/773,371
; PRIOR FILING DATE: 2001-02-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 723
; TYPE: PRT
; ORGANISM: Human
US-10-274-878-2

Query Match 100.0%; Score 3774; DB 4; Length 723;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 723; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MILBERPDGAGAGEEPRLQISRRKPRKTRVSSLRGREGRLDVCAMRGCRVHRGEDPVR 60
Db 1 MILBERPDGAGAGEEPRLQISRRKPRKTRVSSLRGREGRLDVCAMRGCRVHRGEDPVR 60
QY 61 VHVGMNPNQLHAGVCDSTQTCGQLOQSRRAQIHQOIQDKELQMTGAEENLRYATSNRVR 120
Db 61 VHVGMNPNQLHAGVCDSTQTCGQLOQSRRAQIHQOIQDKELQMTGAEENLRYATSNRVR 120
QY 121 ETVALELSYVNSNLQLLKEELEELSGGVDVDPGHGSEAVTVPMIPLGKETKELDWSPLK 180
Db 121 ETVALELSYVNSNLQLLKEELEELSGGVDVDPGHGSEAVTVPMIPLGKETKELDWSPLK 180
QY 181 ELISVHFGEADGASYAEIREALEALQAMRTPSRNBSGLELLTAYYNQICFLDARELTPAR 240
Db 181 ELISVHFGEADGASYAEIREALEALQAMRTPSRNBSGLELLTAYYNQICFLDARELTPAR 240
QY 241 SLGLFFHWYDSLTVGPAQORALAFKGSVLFNIGALHTQIGARQDRSCTEGARRAMEAFQ 300

Db 241 SLGLFFHWYDSLTVGPAQORALAFKGSVLFNIGALHTQIGARQDRSCTEGARRAMEAFQ 300
QY 301 RAAGAFSLLENFSPAPSDMSAASICALPQLMMAQAQECVFEGLSPASMAPODCLAQL 360
Db 301 RAAGAFSLLENFSPAPSDMSAASICALPQLMMAQAQECVFEGLSPASMAPODCLAQL 360
QY 361 RLAEAAQVAAEYRLVHRTWAQPPVHDYVPVSWTALVHVKAAYFRSLAHYHVAMALCDGS 420
Db 361 RLAEAAQVAAEYRLVHRTWAQPPVHDYVPVSWTALVHVKAAYFRSLAHYHVAMALCDGS 420
QY 421 PATEGELPTEHQVFLQPTTSKPRGPVLPOELERRQLGKHLKRAILGQEEALRLHALC 480
Db 421 PATEGELPTEHQVFLQPTTSKPRGPVLPOELERRQLGKHLKRAILGQEEALRLHALC 480
QY 481 RVLRVDLLRAVISQTLQSLAKYAELEDRDDFCEAAEAPDIQPKTHQKPEARNPRLSQG 540
Db 481 RVLRVDLLRAVISQTLQSLAKYAELEDRDDFCEAAEAPDIQPKTHQKPEARNPRLSQG 540
QY 541 KGPDIHRLGFLSVFSAKNRWLVGPHVLTGRGGFGLTLRGDSPVLIAAVIPGSQAAAA 600
Db 541 KGPDIHRLGFLSVFSAKNRWLVGPHVLTGRGGFGLTLRGDSPVLIAAVIPGSQAAAA 600
QY 601 GLKEGDYIVSVNGQPCRWWRHAEVVTTELKAAGEAGASLOVVSLLPSSRLSLGDRRPVLL 660
Db 601 GLKEGDYIVSVNGQPCRWWRHAEVVTTELKAAGEAGASLOVVSLLPSSRLSLGDRRPVLL 660
QY 661 GPRGLLRQREHGCKTTPASTWASPRLLNWSRKAQCGKTCGCPQPCAPVKPAPPSLKH 720
Db 661 GPRGLLRQREHGCKTTPASTWASPRLLNWSRKAQCGKTCGCPQPCAPVKPAPPSLKH 720
QY 721 GWP 723
Db 721 GWP 723

RESULT 3
US-09-849-334-4
; Sequence 4, Application US/09849334
; Patent No. 6500655
; GENERAL INFORMATION:
; APPLICANT: RUSCH, Douglas et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001099-CIP
; CURRENT APPLICATION NUMBER: US/09/849,334
; CURRENT FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 634
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-849-334-4

Query Match 61.8%; Score 2331; DB 4; Length 634;
Best Local Similarity 67.0%; Pred. No. 3.3e-210;
Matches 473; Conservative 49; Mismatches 112; Indels 72; Gaps 6;
QY 1 MILBERPDGAGAGEEPRLQISRRKPRKTRVSSLRGREGRLDVCAMRGCRVHRGEDPVR 60
Db 1 MILBERPDGAGAGEEPRLQISRRKPRKTRVSSLRGREGRLDVCAMRGCRVHRGEDPVR 60
QY 61 VHVGMNPNQLHAGVCDSTQTCGQLOQSRRAQIHQOIQDKELQMTGAEENLRYATSNRVR 120
Db 27 -----KGYGSPVQNPQGLQSHRARLHQOISKELMRMRTGAENLRYATSNRVR 74
QY 121 ETVALELSYVNSNLQLLKEELEELSGGVDVDPGHGSEAVTVPMIPLGKETKELDWSPLK 180
Db 75 ETVALELSYVNSNLQLLKEELEELSGGVDVDPGHGSEAVTVPMIPLGKETKELDWSPLK 134
QY 181 ELISVHFGEADGASYAEIREALEALQAMRTPSRNBSGLELLTAYYNQICFLDARELTPAR 240

Db 135 ELISHFGEDGTSFETEIQELEDLRQATRTPSRDEAGLLDAAAYSQLCLDARFFSPSR 194
QY 241 SLGLFFHYDLSLTGVPAQORALAFKSGSVLFNIGALHTQIGARQDRSCTEGARRAMEAFQ 300
Db 195 SPGLLFHYDLSLTGVPAQORALAFKSGSVLFNIGALHTQIGARQDRSCTEGARRAMEAFQ 254
QY 301 RAAGAFSLLENFNSHAPSPDMSAASLCALEQLMMAQACEVFEGLSPASMAPODCLAQL 360
Db 255 RAAGAFSLLENFNSHAPSPDMSAASLCALEQLMMAQACEVFEGLSPASMAPODCLAQL 314
QY 361 RLAEAAQAAVEAYEGLVHRMAQPPVRDYLPAWNTNLAHVKAHFCALAHYHAAALCESH 420
Db 315 QLAQEAQAAVEAYEGLVHRMAQPPVRDYLPAWNTNLAHVKAHFCALAHYHAAALCESH 374
QY 421 PATEGELPTHQVFLQPTSSKPRGVPVLPQELERRQLGKAHLKRAILGQBEALRLHALC 480
Db 375 PA-KGELARQEHVF-QPSTPHEPLGPTLPQHPEDRRKLAKAHLKRAILGQBEALRLHALC 432
QY 481 RVLREVDLLRAVISOTLQSLAKYAELEDDDFCEAAEAPDIQPKTHOKPEARMPRLSQG 540
Db 433 RVLKRVLDLQVVVQALRRSLAKYQLEREDDFEATEAPDIQPKTHQTPPE----- 483
QY 541 KGPDIHRLGSLVFSAKNRWLVGPVHLTRGEGFGTLRGDSPLVIAAIVPGSQAAAA 600
Db 484 -----GPLSVFSTKNRQLVGPVHMTREGGFGTLRGDSPLVIAAIVPGSQAAAA 534
QY 601 GLKGGDIYVSVNGQPCRWWRHAEVVTTELKAAAGEAGASLQVSLPSSRLPSIGDRRPVLL 660
Db 535 GLKGGDIYVSVNGQPCRWWRHAEVVTTELKAAAGEAGASLQVSLPSSRLPSIGDRRPVLL 594
QY 661 GPRGLLRSOREHGCKTPASTWASPRPLNWSKKAQOGKGTGGCPOPC 706
Db 595 W-----NQREGFETPMTTRTPWPLGWSKKNKQGTGSHPDPC 634

RESULT 4
US-10-274-878-4
; Sequence 4, Application US/10274878
; Patent No. 6670163
; GENERAL INFORMATION:
; APPLICANT: RUSCH, Douglas et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CLO01099-CIP-DIV
; CURRENT APPLICATION NUMBER: US/10/274,878
; CURRENT FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: 09/849,334
; PRIOR FILING DATE: 2001-05-07
; PRIOR APPLICATION NUMBER: 09/773,371
; PRIOR FILING DATE: 2001-02-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 634
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-274-878-4

Query Match 61.8%; Score 2331; DB 4; Length 634;
Best Local Similarity 67.0%; Pred. No. 3.3e-210;
Matches 473; Conservative 49; Mismatches 112; Indels 72; Gaps 6;
QY 1 MILERPPGAGAGESPRLQISRRKPRKTRVSSLRGRREGLDVCAWRCGRVHRGEDPVR 60
Db 1 MILERPPGQGTGESSRPPQD-----GSIR----- 26
QY 61 VRVGPWNLHAGVCDLSLTQCGLOSRRAQHQQIDKELOMRTCAENLYRATSNRVR 120
Db 27 -----KGVGSFVQNPQGLQSHRAKLHQQISKELSMRGTGAENLYRATSNRVR 74
QY 121 ETVALELSYVNSNLQLLKEELAEELSGVDPPGRHGSEAVTVPMIPGLKTKETKELDWSTPLK 180

Db 75 ETVALELSYVNSNLQLLKEELAEELSTSVDDVQPEGEGTIPMIPLGLKETKELDWSTPLK 134
QY 181 ELISHFGEDGASYEAEIRELEALRQAMRTPSRNEGLELLTAYYNQLCFDLARELTIPAR 240
Db 135 ELISHFGEDGTSFETEIQELEDLRQATRTPSRDEAGLLDAAAYSQLCLDARFFSPSR 194
QY 241 SLGLFFHYDLSLTGVPAQORALAFKSGSVLFNIGALHTQIGARQDRSCTEGARRAMEAFQ 300
Db 195 SPGLLFHYDLSLTGVPAQORALAFKSGSVLFNIGALHTQIGARQDRSCTEGARRAMEAFQ 254
QY 301 RAAGAFSLLENFNSHAPSPDMSAASLCALEQLMMAQACEVFEGLSPASMAPODCLAQL 360
Db 255 RAAGAFSLLENFNSHAPSPDMSAASLCALEQLMMAQACEVFEGLSPASMAPODCLAQL 314
QY 361 RLAEAAQAAVEAYEGLVHRMAQPPVRDYLPAWNTNLAHVKAHFCALAHYHAAALCESH 420
Db 315 QLAQEAQAAVEAYEGLVHRMAQPPVRDYLPAWNTNLAHVKAHFCALAHYHAAALCESH 374
QY 421 PATEGELPTHQVFLQPTSSKPRGVPVLPQELERRQLGKAHLKRAILGQBEALRLHALC 480
Db 375 PA-KGELARQEHVF-QPSTPHEPLGPTLPQHPEDRRKLAKAHLKRAILGQBEALRLHALC 432
QY 481 RVLREVDLLRAVISOTLQSLAKYAELEDDDFCEAAEAPDIQPKTHOKPEARMPRLSQG 540
Db 433 RVLKRVLDLQVVVQALRRSLAKYQLEREDDFEATEAPDIQPKTHQTPPE----- 483
QY 541 KGPDIHRLGSLVFSAKNRWLVGPVHLTRGEGFGTLRGDSPLVIAAIVPGSQAAAA 600
Db 484 -----GPLSVFSTKNRQLVGPVHMTREGGFGTLRGDSPLVIAAIVPGSQAAAA 534
QY 601 GLKGGDIYVSVNGQPCRWWRHAEVVTTELKAAAGEAGASLQVSLPSSRLPSIGDRRPVLL 660
Db 535 GLKGGDIYVSVNGQPCRWWRHAEVVTTELKAAAGEAGASLQVSLPSSRLPSIGDRRPVLL 594
QY 661 GPRGLLRSOREHGCKTPASTWASPRPLNWSKKAQOGKGTGGCPOPC 706
Db 595 W-----NQREGFETPMTTRTPWPLGWSKKNKQGTGSHPDPC 634

RESULT 5
US-09-800-729-199
; Sequence 199, Application US/09800729
; Patent No. 6605592
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: 32 Human secreted proteins
; FILE REFERENCE: P2044P1
; CURRENT APPLICATION NUMBER: US/09/800,729
; CURRENT FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: PCT/US00/26013
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 60/155,709
; PRIOR FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 199
; LENGTH: 921
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (247)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (362)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (603)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-800-729-199

Query Match 7.1%; Score 268; DB 4; Length 921;
Best Local Similarity 26.7%; Pred. No. 6.4e-16;

Db 114 ----GAGGTCSATNLSRVAGLEKQLAIELKVKQGAENMIQTYSNGSTKDRKLLLTAAQM 169
QY 243 GUFFHYDLSLT-----GVPAQORALAFKGVSVLNTGALHTQIGA-----282
Db 170 ----LQDSKTKIDIIRMQLRALQAGLENAAPDDTQOSP--DLGAVELRIEELRHIF 222
QY 283 RODRSCTEGARAMEAFORAAGAFSLLENFNSHAPSPDMSAASLCALEQLMMAQAQECVF 342
Db 223 RVEHAVAEGAKVRL-----LSNAKAPDRKAVSEAEQKUTESQKLGILLR 268
QY 343 EGLSPSPASMAPQCLQLRACAAQVAAEYRLVHRTMAQPPVHDY-----VPVSWTALVH 398
Db 269 EALERRLGLPAD-HPKGLRLLELAASAASAFSTRLAGFPFATHYSTLCKPAPLTGTL 327
QY 399 VKAEYFRSLAH-----YHVAVALCDGSPATEGELPTHEQVFLQPP-----438
Db 328 VRVVGCRDLPETIPWNPFTSM--GGFGT----PDSRPFLSRPARGLYSRSGSLSGRSSL 381
QY 439 ----TSSKPRGP-----VLPQLEERQLGKAHLKRAILGQ 470
Db 382 KAAEANTSEVSVLKLNDTVVGOTSWKPCGNWDQSFTLELERAKE-----LELAVFWR 436
QY 471 BEALRLHALC--RVLEVDLLRAVISQTLQRSLSAKYAELEDREDDFCEAAEAPDIOPKTHQ 528
Db 437 DQ----RGLCALFKLEDFL-----DNERHEVQDMREFQGCCLVAEV-----TFR 477
QY 529 KPE-ARMPEL-----SQKGPDIFHRLGPLSVFSAKNRWLVGPHLTFR-----571
Db 478 NVIEIRIPRLRQKIFSKQOGKA--FORAQMNIDVA--TW-----VRLLERLIPNAT 527
QY 572 GEGGGLTLRGDSPVLIAAIPGSOAAAAGLKEGDYIVSVNGQPCRWHRHAEVVTTELKAA 631
Db 528 GNGTF-----SP-----CASPGSEARTTG-----DISVE-----KING 556
QY 632 GEAGASLQVSVLLPSRLPSLDGRPRVLGPRGLLRSOREHGCKTPASTWASPRLLNWS 691
Db 557 TDSDSPQKSSRDPSPSSLS--PI-----QBSTAPELPSETQETPGPAL----601
QY 692 RKAQOGKGGCPQCAPVKPAP 713
Db 602 -----CSPLRKSP 609

RESULT 8

US-08-665-037-2
; Sequence 2, Application US/08665037
; Patent No. 5895813

; GENERAL INFORMATION:
; APPLICANT: Seedorf, Klaus
; APPLICANT: Ullrich, Axel
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT
; TITLE OF INVENTION: OF TKA-1 RELATED
; TITLE OF INVENTION: DISORDERS
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/665,037
; FILING DATE: June 13, 1996
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 60/005,167
; FILING DATE: October 13, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 220/156
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 450 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-665-037-2

Query Match 3.6%; Score 134.5; DB 2; Length 450;
Best Local Similarity 24.6%; Pred. NO. 0.00071;
Matches 84; Conservative 21; Mismatches 111; Indels 125; Gaps 14;

QY 486 VLLRAVISQTLQRSLSAKYAELED-----REDDFCEAAEAPDIOPKTHQKPEARMPRLSQCK 541
Db 72 VQRIKAVEGQT--RLLVVDQETDEELRRQLCTCTEMAQORGLPPADHPWEPK-----121
QY 542 GPDIFHRLGPLSVFSAKNRWLVGPV-----HLTRGEGGGLTIRGDSPL--VLIAA 590
Db 122 -FDWAHTGSHSEAGKGD--VSGPLRELRLPRLCHLRKGPQGVGNLHSDKSRPGQYIRS 177
QY 591 VTPGQAAAAGLKEGDYIVSVNGQPCRWHRHAEVVTTELKAAEAGASLQVV-----641
Db 178 VDPGFGARGSLRAQDRLIEVNGQNVGLRHAEEVVASIKAR-EDEARLLVDPETDEHEFK 236
QY 642 -----SLLPS-----SRLP-----SLGDRPP-----657
Db 237 RLVRVPTTEHVEGFLPSPYNTGTSAPQLNGGSACSRSDLPGSDKDTEDGSAWKQDPQOE 296
QY 658 --VLLGPRGLLRSOREHGCKT-----PASTWAS 683
Db 297 SGLHLSPTAAEAARRRLEPCESTSARHRTGTGTSVKSSATSEPLPACLGTLGLPLHGPWAS 356
QY 684 PRPLNWSRKAQOGKGGCPQCAPVKPA---PPSSLKHGP 721
Db 357 ACPEL-----PQOWTGMSCHEPCPEISPGEPSPCPGPP 392

RESULT 9

US-08-666-067-2
; Sequence 2, Application US/08666067
; Patent No. 5922842
; GENERAL INFORMATION:
; APPLICANT: Seedorf, Klaus
; APPLICANT: Ullrich, Axel
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT
; TITLE OF INVENTION: OF TKA-1 RELATED
; TITLE OF INVENTION: DISORDERS
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:

```
; APPLICATION NUMBER: US/08/666,067
; FILING DATE: June 13, 1996
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/005,421
; FILING DATE: October 13, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 220/157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 450 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-666-067-2

Query Match 3.6%; Score 134.5; DB 2; Length 450;
Best Local Similarity 24.6%; Pred. No. 0.00071;
Matches 84; Conservative 21; Mismatches 111; Indels 125; Gaps 14;

QY 486 VDLRAVISQTLQSLAKYAEID---REDDFCEAAEAPDIQPKTHQKPEARMPLRSQK 541
Db 72 VQRIKAVEGT--RLVVVDQETDEELRRQLCTTEEMAQRGLPPAHDPWEK----- 121
QY 542 GPDIFHRLGPLSVFSAKNRMVLGVF-----HLTRGEGGGLTLRGDSP---VLIAA 590
Db 122 -PDWAHTGSHSSEAGKXD---VSGFLRELRLPRLCHLRKPGQYGFNLHSDKSRPQYIRS 177
QY 591 VIPGQAAAAGLKEGDYIVSNGQPCRWRAEAVVTELKAAGEAGASIQVY----- 641
Db 178 VDPGSPGARSGLRAQDRLLIEVNGQNVGLRHAENVASIKAR-EDEARLLVVDPTDEHFK 236
QY 642 -----SLLPS-----SRLP-----SLGDRRP--- 657
Db 237 RLRVTPTTEHVEGFLPSPVNTGTPAQLNGSGSACSSRDLFGSKDXTEDGSAWKQDPQ 296
QY 658 --VLLGPRGLLRSOREHGCKT-----PASTWAS 683
Db 297 SGLHLSPTAAEARRRLPECESTSARHRTGTGSKSSATSBELPACLTGLGLPHGPWAS 356
QY 684 PRPLNWSRKAQCKTGCCPQPCAPVKA---PPSSLKHG 721
Db 357 ACPEL-----PQOWTGWGWSCHCEISPSGPEPPSCPCPPG 392

RESULT 10
US-08-732-870-2
; Sequence 2, Application US/08732870
; Patent No. 5945523
; GENERAL INFORMATION:
; APPLICANT: Ullrich, Axel
; APPLICANT: Seedorf, Luitgard
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT
; TITLE OF INVENTION: OF TKA-1 RELATED
; TITLE OF INVENTION: DISORDERS
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESS: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
```

```
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/732,870
; FILING DATE: October 15, 1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/666,037
; FILING DATE: June 13, 1996
; APPLICATION NUMBER: 08/666,067
; FILING DATE: June 13, 1996
; APPLICATION NUMBER: 60/005,167
; FILING DATE: October 13, 1995
; APPLICATION NUMBER: 60/005,421
; FILING DATE: October 13, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 222/247
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 450 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-732-870-2

Query Match 3.6%; Score 134.5; DB 2; Length 450;
Best Local Similarity 24.6%; Pred. No. 0.00071;
Matches 84; Conservative 21; Mismatches 111; Indels 125; Gaps 14;

QY 486 VDLRAVISQTLQSLAKYAEID---REDDFCEAAEAPDIQPKTHQKPEARMPLRSQK 541
Db 72 VQRIKAVEGT--RLVVVDQETDEELRRQLCTTEEMAQRGLPPAHDPWEK----- 121
QY 542 GPDIFHRLGPLSVFSAKNRMVLGVF-----HLTRGEGGGLTLRGDSP---VLIAA 590
Db 122 -PDWAHTGSHSSEAGKXD---VSGFLRELRLPRLCHLRKPGQYGFNLHSDKSRPQYIRS 177
QY 591 VIPGQAAAAGLKEGDYIVSNGQPCRWRAEAVVTELKAAGEAGASIQVY----- 641
Db 178 VDPGSPGARSGLRAQDRLLIEVNGQNVGLRHAENVASIKAR-EDEARLLVVDPTDEHFK 236
QY 642 -----SLLPS-----SRLP-----SLGDRRP--- 657
Db 237 RLRVTPTTEHVEGFLPSPVNTGTPAQLNGSGSACSSRDLFGSKDXTEDGSAWKQDPQ 296
QY 658 --VLLGPRGLLRSOREHGCKT-----PASTWAS 683
Db 297 SGLHLSPTAAEARRRLPECESTSARHRTGTGSKSSATSBELPACLTGLGLPHGPWAS 356
QY 684 PRPLNWSRKAQCKTGCCPQPCAPVKA---PPSSLKHG 721
Db 357 ACPEL-----PQOWTGWGWSCHCEISPSGPEPPSCPCPPG 392

RESULT 11
US-09-252-991A-30851
; Sequence 30851, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
```


;
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30851
; LENGTH: 1266
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30851

Query Match 3.6%; Score 134; DB 4; Length 1266;
Best Local Similarity 22.1%; Pred. No. 0.0043;
Matches 178; Conservative 81; Mismatches 288; Indels 260; Gaps 44;

QY 3 LRRPDGAGAGEEPLOISRRKPRKTRVSSLR-GRREGLDVCAWRCRVHGEDPVRV 61
DB 421 IQRRELGEPPGPHGLQPGAR--RRVRAGMKRQRREHLAD-----HRGED--QR 467
QY 62 HVGPMNPQLHAVGCDLSLTQICGQQLQSRRAQIHQIDKELQMTGAENLYR----- 112
DB 468 HGGRO-----GC--LRPRAGLDLRR--VHQRREGGQGGAGRRRRHHGDAGAGL 516
QY 113 ---ATSNRVRTVALELSYVNSNLIQLKELEELSGGVDPRHSGSEAVTVPMLPLGLKE 169
DB 517 FLQAARDRAFPQGRQHRPAGDGLQ---QPADLPERRHPGH-----FGFPGLRKE 564
QY 170 TKELDWSTPLKELISVHFCEDGASYEABIRELEALRQAMRT-PSRNESGLELLTAYYNOL 228
DB 565 HRLLOGQ--LRHPPLHRRTPQGG-----RALHPLRARFRRGFGKRGGG----- 606
QY 229 CFLDARFLTPARSLGFLFFHWDLSLTGVPAQR-----ALAFKGSVLFNT----- 273
DB 607 -----CPGLDLG-----NVQVPEGRDRDHLPPGGRALRGDADLRVADDDPSR 651
QY 274 -----GALHTQIGA-----RODRCTEGARPAEAFQAAAGAFSILLRENFHSHAPSDMS 322
DB 652 RPSREGVAGHQAALRATRRSRQRPHPSTAPGARRRGRARRA-----DHGQGPQP 700
QY 323 AASLCALQQLMAQAQECVFEGLSPPAS-----MAPQDCLAQLRLAQEAQVAABYRLVH 377
DB 701 AG--AAHRLVADAPP-----SPRSPGSGRGP-----AABHS-VH 734
QY 378 RTMAQPPV---HDYVPVSWTA-----LVHVKAIEYFRSLAHYHVAMALCDGSGPATEGELP 428
DB 735 RPQGGPMPSILGHNVGGARSAGNLIILSLDSDSSEALPYAFV-----QATEAEV- 785
QY 429 THEQVLPQPTSSKPRGPVLP--QELERRQLGKAHLKRAILGQBEALRLHALCRVLRV 486
DB 786 -----DAAAARAAERAYPHYRLSATRRAG---FLEATASRLDALGDFFVALVRRET 833
QY 487 DLLRAVISQTLQRS---LAKYAEALDREDDCEAAEAPDIQ--KTHQKPEARMPLRSQKG 542
DB 834 ALPAARIQERTRTANQURLFAEVLRRGDF-HGAIDRGQPGORTPPRPLRQWRIG--- 888
QY 543 PDIFHRLGPLSVFSKAKNRWLVPVHLTRGEGGFLTLRGDSPVLIAAVIPGSAQAAGL 602
DB 889 -----LGPVAVFGASN-----FPLAFSTAGDSAAALAAAGCPVVXA----- 925
QY 603 KEGDYIVSVNGQPCRWVHAEVWTE--LKAGEAGASLOVSLSPSSRLPSLGRRPVL- 659
DB 926 -HGGMAT-----AECVADAILQAAADSGMPAGVFNMYVYSGVGGEALVRHPAIR 973
QY 660 -LGRPGLLRSOREHGCKTPASTWASPRLLNWS-----RAAQ----- 696
DB 974 AVGFTGSLKGRAL-CDLAA---ARPQIPVFAEMSSINPLVLPEALRRGRQVAEELA 1029
QY 697 -GKTGGCPQCAPKVPAPFPSSLKHPGW 722
DB 1030 ASVTILGCGQFC--TKPGLVLGRSPGF 1054

RESULT 12

US-09-368-590-2

;
; Sequence 2, Application US/09368590
; Patent No. 6187563
; GENERAL INFORMATION:
; APPLICANT: Solimena, Michele
; TITLE OF INVENTION: INTERACTING POLYPEPTIDES FOR
; TITLE OF INVENTION: AUTOCANTIGENS OF AUTOIMMUNE DISEASES
; FILE REFERENCE: 101918-200 (OCR-941)
; CURRENT APPLICATION NUMBER: US/09/368,590
; CURRENT FILING DATE: 1999-08-04
; EARLIER APPLICATION NUMBER: 60/095,657
; EARLIER FILING DATE: 1998-08-07
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 2293
; TYPE: PRT
; ORGANISM: Human
US-09-368-590-2

Query Match 3.5%; Score 131.5; DB 3; Length 2293;
Best Local Similarity 21.2%; Pred. No. 0.019;
Matches 161; Conservative 102; Mismatches 296; Indels 201; Gaps 38;

QY 29 TRVSSLRGEREGLDVCAWRCRVHGEDPVRHVGPMNPQLHAVGCDLSLTQICGQLOS 88
DB 1529 TAAATMAEKMGLINE--AWAELELMG---TRAQLLAASRELHKFFSDA-RELQ-GQIEE 1581
QY 89 RRAQHQQIDKELQMTGAENLYRATSNRVRTVALELSYVNSNLIQLKELEELSGG- 147
DB 1582 KERRLP-----RLTTPPEPSPASSMQRT-----LRAFEHLQLLVSVQRLQEGA 1627
QY 148 -----VDPRHSGSEAVTVPMLPLGKTKELDWSTPLKELIS-----VHFGEDGASY 194
DB 1628 AQLRVTYAGEH-AEAIA-----SREQEVLQGWKELLSACEDARLHVSTADALRF 1676
QY 195 EABIRELEALRQAMRT-----PSRNESGLELLTAYYNOL-CFLDARF--LTPARSLGLF 245
DB 1677 HSQVRDLLSWMDGIASQIGAADKPRDVSSEVLMNYHQGLKTELEARVPELTTCQELG-- 1734
QY 246 FHWYDLSLTGVPAAQRALAFKGSVLFNIGALHTQIGAPQDRSCTEGARRAMEAFQARA 305
DB 1735 -----RSILLINKSMADEITQAKLDKLTKEEVSERKDRH-MEWLQOMLEHVHQAQEA 1786
QY 306 FSLIRENFHSHAPSDMSA-----ASLCALQQLMAQAQECVFEGLSPPASMAPODCLAQLR 361
DB 1787 --VVADAMLTAEQPLQLQSRRELGSVDEVEQLIRRH--AFR---KAAAWEERFSSIR 1837
QY 362 LAQEAQAQVAABYRLVHRTMAQPPVHDYVPVSWTALVHVKAIEYFRSLAHYHVAMALCDGSP 421
DB 1838 RLTTIEKIAE-----QSKQPP-----TPLLGRK--FF-----GDP 1866
QY 422 ATEGELPHEQVFLQPPTSSKPRGPV-----LPOELERRQLGKAHL 463
DB 1867 T---ELAAKAAPLRPGYERGLFELARRASDTLSAEVTRTVGVYRQELKEERLQPRIDR 1923
QY 464 KRAILGQBEALRLHALCRVLRREVLLRAVISQT-----LQSLAKYAEALDREDDFCEAAE 518
DB 1924 LPEIPGRVEPAALPAAPEDAAETATPAABQVPRPERQESADRAEELPR-----R 1976
QY 519 APDIQPTXTHQKPEA---RMPRLSQKGPDIFHRLGPLSVFSKAKNRWLVPVHLTRGEGG 575
DB 1977 RPEQESVDQSEEAARRRRRPERQESAHEAAHSL--TLGRYEQMERRRRRRERLERQESS 2035
QY 576 -FGTLRGDSPVLIAAVIPGSAQAAGLKEGDYIVSVNGQPCRWVHAEVWTE--KAA 632
DB 2036 EQEMPIRGD-----LVKGAHL-----ADIVEQLQEKAG 2065
QY 633 ---EAGASLOVSLSPSSRLPS---LGDRRPVLGPRGLLRSQREHGCKTPASTWASPRP 686
DB 2066 PGLPAGPSLPQPRELPPGLPGLNGLELPEPTPRPDRPRARORPKPR-----RPRP 2115
QY 687 LL-----NWSKAAQCGKTCGCPQPCAPVKPAPPSLKHPGW 722

Db 2116 RECCEGGRSRSAQAQGSAPP--PPPTHVQHEGF 2153

RESULT 13
US-09-252-991A-27424
; Sequence 27424, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27424
; LENGTH: 751
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27424

Query Match 3.4%; Score 129; DB 4; Length 751;
Best Local Similarity 21.9%; Pred. No. 0.0054;
Matches 185; Conservative 80; Mismatches 254; Indels 326; Gaps 47;

Qy 28 KTRVSSLRGR---REGLRDVCAMRGCRVHRGDEPVRVH-----VGMNPNQLHVGCD 76
Db 33 KARLDSQERSQASSLRDLG-----DAIRLERELADMRKVGDRGVAEHA----- 78

Qy 77 SLTQCCQSQSRAQHQIDKELQMTGAENLYRATSNRVRRETVALELSYNSNQL 136
Db 79 -----RQLGEROQL-----KRLGIE--ARAAGDAYAR----- 104

Qy 137 LKEELERLSGCVDPGRHSGEAVTPMPLGKTKETKELDWSTPLKELIS-----VHFGEDGAS 193
Db 105 -----LGMQORGLD-----MQVRGLQRLEQASQAMPLASAFGLVVEASKTAAG 148

Qy 194 YEATIRELEALRQAM---RTP-----SRNESGLELLTAYYNOLCFDAR---FLPPA 239
Db 149 YQARLRDL--ALRGLDVGREPALASLIQDSANOSGLG--RTATLDMLEHLNATGMGFAAAQ 206

Qy 240 RSLGLF--FHWYDSLITG--VPAQORALAFKGSVLFNIGALHTQIGARQDRSCTEG--ARR 294
Db 207 MNMGLAGRFQGGIASAFAVAGLVRLALQAQGS-----DSPEQLSASLDRLVLKGRV 260

Qy 295 AMEAFQRAAGAFSLLRENFSHAPSMDMSAASLCALQMLMAQAQRCV----- 341
Db 261 GSEALARLRLPALLSALGNAGEATAGDVG-----ALGALLEIOAKNTTPDKADVRMKAWLE 315

Qy 342 FEGLSPPSMAPODC---LAQIR-----LAQAAQVAAEYR-----LVHRTMAQPPVHDYV 389
Db 316 FVGSGLKRAYQDYDRDLALRKDGASLLEANLELAARYDKGKLSAGVASPALEAY- 374

Qy 390 PYSWTALVHVKAIEYFRSLAHYHVAMALCDGSPATEGELPTHQVFLQPTTSS-----K 442
Db 375 -----RASGEF-----QGLLESQSSVSGSERDA 399

Qy 443 PRGPVLPQBL-----ERRQ--LGKA-----HLKR--AILGOEALRLHALCRVLR- 484
Db 400 QRRKMSQELWKASSDSWERAQTALGSALNPYLDNLAKGSVILGESTABLEAYPRTTAG 459

Qy 485 -----EVDLLR-----AVISQTLQSLAKY---AELDR--- 509
Db 460 LTAAGAVLSGLYAKGREGAIDVLRGRLGRGTAAVGLDTERGAGRVSGSGSEIQRVPV 519

Qy 510 -----EDFCEAAEPDIPKTHQKPEARMPLSQGKGPDIHRLGLPLSVFSKRNWR 562
Db 520 TNWVPVGGDSTLESARRP-----AQRKRGQTPRRKRGKGGGLKARSLSLGSFA----- 568

Qy 563 LVGPVHLTRGEGGFG-----LTLRGDSPVLIA-AVIFGSQAAAAGLKEGDYIVS 610
Db 569 -----GGGLGAMAGKLPRLSRLPTR-NAPLQVASSLIDVAEYVSDLSSEKTI- 616

Qy 611 VNGQPCRWHEAEVVTTELKAAGAGASLQVSVLLPSSRLPSLGDRRPV---LLGPRGLLR 667
Db 617 -----AYGEAGASL-AGSLAGALGASIGSVVPPVVGTLIG--GLVG 654

Qy 668 SQREHGCKTPASTWASPRPLNWSRKAQCKTG---GCP-----OPCAPVKEPAPSS 717
Db 655 G-----AICANGGSELGRLGSLAGDPPAASDNKPAVPVPOAEFVAA 697

Qy 718 KHPCW 722
Db 698 V-PNW 701

RESULT 14
US-09-252-991A-18483
; Sequence 18483, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18483
; LENGTH: 658
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18483

Query Match 3.4%; Score 128.5; DB 4; Length 658;
Best Local Similarity 20.9%; Pred. No. 0.0048;
Matches 158; Conservative 84; Mismatches 300; Indels 215; Gaps 33;

Qy 7 PDGAGAGES-----PRLQ--ISRRK---PRKTRV-----SSLRGRREGRLDVCANRG 49
Db 15 PPAATAGDRSSVEDPFRHPLHQAALSQRNVLPEQHRAILPEFVGLQGRQ-LAVLVALAG 73

Qy 50 CRVHRGEPVVRVHVGMNPNPOLHAVGCDSLTQICQQLQSRRAQJHQOIKELQ--MRTG- 106
Db 74 RRIQGGEDAAGTGLG--HDELDRADADAPPV---LGERGAAFHQVGTGAQHLQRPAD 127

Qy 107 --AENLYRATSNRVRRTVA---LELSYVNSNLQLLKEELBELSGVDPGRHGSSEAVTP 161
Db 128 PLAENVVERSLADQOQRETIVGEMHLLPRFVAGIDLPRLRL---GNVHQARRAAEERQ 183

Qy 162 MIPGLKETKELDWSTPLKELISVHFG---EDGASYAEIREFEALQAMR--TPSRNE 215
Db 184 -VRCGVAGIRRLQGIAPDRDATQVGVGCRYFEAQGHPRRIQPDRLRRPARKLDTFQLRQ 242

Qy 216 SGLLELLTAYYNQLCFDARFLTPARSLGLFHFHWYDSLTVGPAQORALAFKGSVLFNIGA 275
Db 243 OALGETS-----GFGVHRHPASVADHAQR-----IERDA 272

Qy 276 LHTQIGARQDRSCTEGARRAMEAFORAAGAFSLLKENFSHAPSMDMSAASLCALQMLMA 335
Db 273 LQGLGGRLDRITAVQ--RGETEHGTSILGGKMLL---YGVAGKADKTOTPLQAADQ--- 323

Qy 336 QAQECVFEGSLSPASWAPQDCLAQRLACERAAQVAAEYVELVHRTMAQPPVHDYVPSWTA 395
Db 324 -----SAPPSV-----LARRLAQQRRTGQVHAQARAGNRRETQLASFQRL 363

Qy 396 LVHVKAIEYFRSL- - - - -AHYHVAMALCDGSPATEGELPTHQVFLQ----- 436

Db 364 AVEIDVETFRRLHFQWHPRADEAQSHTVATAB-QGAQGTDDHVLRRRQDAHVQQAUVVHSR 422
QY 437 -----PPTSSKRGVPLPOELE-ERRQLGKAHL-KRAILGQBEALRLHALCRV 482
Db 423 LRTQYVPAAGLAPVTDQOGEELAFPLEVGAETALGMEHFGKTGQPGQVGGTGAENLLQV 482
QY 483 LREV--DLLRAVISQTLQSLAKY-AELDREDDFCEAAE-----APDIQPKTH 527
Db 483 LRGIIRDLEATAASGHVEQLAVHLAQVDRPRELVEQRQRGRFQNGPGAGEIVGGAQR 542
QY 528 QKPEARMPLSOGKGPDI FHRIGPLSVFSKAKNRWLVGPHLTRGEGGFLTLRGDSPVL 587
Db 543 QDQRAFLASPARQR-----FCHLAQRAVATAGDQRRTAGRQLL-----DDAPGV 587
QY 588 IAAVTPGSQAAAAGLKEGYIYVSVNGQPCRWHEHAEVVTTELKAAAGEAGASLQVVS 643
Db 588 --AVLPG-----YSHRQRPGLALRRDRGTNLFVARLLAQ 621
QY 644 -----LPSSRLPSLGDRRPVLIG 661
Db 622 DQNLVLVHRTPLVALSLPIDRVSGTEVRFPVQAG 658

RESULT 15

US-09-883-096-2
; Sequence 2, Application US/09883096
; Patent No. 6680369
; GENERAL INFORMATION:
; APPLICANT: Heraud, Christophe
; APPLICANT: Craven, Andrew
; APPLICANT: Yu, Ming
; APPLICANT: Sakowicz, Roman
; APPLICANT: Patel, Umesh A.
; APPLICANT: Davies, Katherine A.
; TITLE OF INVENTION: NOVEL MOTOR PROTEINS AND METHODS FOR THEIR USE
; FILE REFERENCE: 020552-00141005
; CURRENT APPLICATION NUMBER: US/09/883,096
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 09/594,655
; PRIOR FILING DATE: 2000-06-15
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2
; LENGTH: 864
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Amino acid sequence encoded by human kinesin motor
; OTHER INFORMATION: protein gene HsKip3a (Figure 1).
; OTHER INFORMATION: Description of Artificial Sequence: Amino acid
; OTHER INFORMATION: sequence of HsKip3a.
US-09-883-096-2

Query Match 3.4%; Score 128; DB 4; Length 864;
Best Local Similarity 20.1%; Pred. No. 0.0084;
Matches 129; Conservative 70; Mismatches 258; Indels 186; Gaps 24;
QY 211 PSRNESGLELIT-AYYNQLCFDARFLTPASLSGLFFHHYDSLTGVPAAQQALAFKGSV 269
Db 193 PASAEQLLEILTRGNRRNTQHTPDANATSSRSHAIFQIFVKQDRVPLGTQAVQVAKMSL 252
QY 270 LFNIGALHTQICARDRSCTEGA--RRAMEAFORAAGAFSLLENFHSHPSPDMSAASLC 327
Db 253 IDLAGSERASTHAKGEFLREGANINRSLALLINVLNALADAKGRKTHVPYRDSKLTREL- 311
QY 328 ALEQIMMAQAEQVFEGLSPSPASMAPQDCLAQLRLAQEA----- 366
Db 312 -LKDSLGNCRCTVMIAAIS-FSSLTYEDTYNTLYADRAKEIRLSKSNVTSLDCHISQY 369
QY 367 -----AQVAA--EYELVHRTWAQPPVHDY--VPVSWTALVHVKAIFY----- 404
Db 370 ATICQQLQAEVAALRKLQVYEGGQPPQDLPSPKSPGPPPEHLPSPLPPHPPSPQCT 429

QY 405 -----RSLAHYHVAMAL-----CDGSPATEGELPTHE---QVFLQPPTSSK 442
Db 430 PELPAGPRALQBEESLGMQAQVERAMEGNSSDQESPEDEDEGFABEVPTQMPQNTHAL 489
QY 443 PRGPVLPOELEERRQLGKAHLKRAILGQBEALRLHALCRVLRVDLLRA--VISQTLQRS 500
Db 490 PESPL--TLQPKPVVGHFSARELDGDRSKQALKVLCVAQRQYSLLOAANLITPDMITE 547
QY 501 LAKYAELEDREDDFCEAAE-----APDIQPKTHOKPEARMPLRSQKGPDIHRLG 550
Db 548 FETLQQLVQEEKIEPCAEALRTSGLARGAPLAQELCSESIIPVPSPLCPPEPG---YTG 602
QY 551 PLSVFSKAKNRWLVGPHLTRGEGGFLTLRGDSPVLIAAVIPGSQAAAAGLKEGYIYS 610
Db 603 PVTRTMAR--RLSGPLH-----TL-----GIPPGNCTFA----- 630
QY 611 VNGQPCRW-----WRHAEVVTTELKAAAGEAGASLQVVSLLPSRLPSLGDRRP----- 657
Db 631 ---QGRWPMKRRRRPSALEADSPMAKRGTKRQRFCLRRGSLPDTQPSQGPSTP 687
QY 658 -----VLLGPRGLIIRSOREHGCKTP-----ASTWASPR 685
Db 688 KGERASSPCHSPRCVPAIVIKSRVELGPSAM-----QNCSTPLALPTFDLNAITDLSSE 741
QY 686 P-----LLNWSRKAQCGTGGCPQPCAPVKPAPPSSLKHP 720
Db 742 PPSKPSFHECIGWDKIPQER--LSRLDQPFIPRAPVPLFTMKGP 782

Search completed: July 13, 2004, 12:08:40
Job time : 33 secs

This Page Blank (uspto)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 12, 2004, 19:58:24 ; Search time 637.213 Seconds
(without alignments)
16460.437 Million cell updates/sec

Title: US-10-697-266-1
Perfect score: 2469
Sequence: 1 tcggcgcccggtggtgcgg.....aaaaaaaaaaaaaaaaaaaaa 2469

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq_29Jan04.*
1: geneseqn1980s.*
2: geneseqn1990s.*
3: geneseqn2000s.*
4: geneseqn2001as.*
5: geneseqn2001bs.*
6: geneseqn2002s.*
7: geneseqn2003as.*
8: geneseqn2003bs.*
9: geneseqn2003cs.*
10: geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2469	100.0	2469	6	AAL48837 Human pro
2	2105	85.3	2837	7	ABV72511 Nucleotid
3	957	38.8	1013	5	AAS00831 Human cDN
4	955.4	38.7	1013	3	AAC77780 Human can
5	952.8	38.6	3213	7	ACC46698 Human dit
6	524.6	21.2	764	2	AZ15546 Human gen
7	496.4	20.1	2100	5	AAS15828 Human cDN
8	496.4	20.1	3526	5	AAS15827 Human ORF
9	494.8	20.0	2109	6	ABX97183 Human NOV
10	493.2	20.0	2061	6	ABX97185 Human GTP
11	493.2	20.0	3484	6	ABX97185 Human GTP
12	487.8	19.8	3019	4	RAF58361 Human cDN
13	463.6	18.8	2310	6	ABX97182 Human NOV
14	418.2	16.9	2757	4	AAC91349 Human pol
15	379.2	15.4	1671	6	ABK92255 Prostate
16	368	14.9	19025	6	AAL48838 Human pro
17	355	14.4	3335	5	AAS72556 DNA encod
18	301.4	12.2	2869	4	AAH75576 Human tra
19	235.4	9.5	966	5	AAS70406 DNA encod
20	219.4	8.9	331	3	AAC76200 Human ORF
21	206.6	8.4	2168	4	ABL14125 Drosophil
22	203.8	8.3	735	3	AAC75784 Human ORF
23	165	6.7	165	2	AAL4379 Human gen

24	150	6.1	933	5	AAS72551	Aas72551 DNA encod
25	143.8	5.8	1330	5	AAS70407	Aas70407 DNA encod
26	117.6	4.8	756	6	ABQ31046	Abq31046 Oligonucl
27	117.6	4.8	756	6	ABQ31047	Abq31047 Oligonucl
28	108.4	4.4	2925	7	ADA52688	Ada52688 Human cod
29	108.4	4.4	5200	6	AAD40740	Aad40740 Human kin
30	108.2	4.4	5234	3	AAC81224	Aac81224 Human his
31	103.6	4.2	599	4	AAL00243	Aal00243 Human rep
32	97	3.9	802	4	AAS03028	Aas03028 Human rep
33	94.6	3.8	439	8	ACH17102	Ach17102 Human dia
34	93.4	3.8	338	4	AAS57421	Aas57421 Human adu
35	87	3.5	479	6	ABL81142	Ab181142 cDNA #97
36	84	3.4	756	6	ABQ31049	Abq31049 Oligonucl
37	84	3.4	756	6	ABQ31048	Abq31048 Oligonucl
38	72.4	2.9	500	6	ABS71881	Abs71881 Human GTP
39	71	2.9	167	5	AAS15837	Aas15837 Human pro
40	71	2.9	389	5	AAS72552	Aas72552 DNA encod
41	69.4	2.8	167	6	ABS71866	Abs71866 Human GTP
42	68.2	2.8	789	5	AAS72550	Aas72550 DNA encod
43	68	2.8	2469	6	AAL48837	Aal48837 Human pro
44	68	2.8	2837	7	ABV72511	Abv72511 Nucleotid
45	66.2	2.7	492	3	AAC74860	Aac74860 Human ORF

ALIGNMENTS

RESULT 1
AAL48837
ID AAL48837 standard; cDNA; 2469 BP.
XX
AC AAL48837;
XX
DT 24-OCT-2002 (first entry)
XX
DE Human protein kinase N family kinase coding sequence.
XX
KW Human; protein kinase N; kinase; PKN; eye retinoblastoma; bocio tumour;
KW placenta choriocarcinoma; germ cell; leukocyte; uterus tumour; cancer;
KW pre-B cell acute lymphoblastic leukaemia; Wilm's tumour; enzyme;
KW brain anaplastic oligodendroma; uterus endometrial adenocarcinoma;
KW cytostatic; gene therapy; chromosome 8; gene; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT 5'UTR 1..102
FT CDS /*tag= a
FT 103..2274
FT /*tag= b
FT 3'UTR /*product= "kinase"
FT 2275..2469
FT /*tag= c
XX
XX WO200261062-A2.
XX
PD 08-AUG-2002.
XX
XX 29-JAN-2002; 2002WO-US002152.
XX
XX 01-FEB-2001; 2001US-00773371.
XX 07-MAY-2001; 2001US-00849334.
XX (PEKE) PE CORP NY.
XX
XX Rusch D, Ketchum KA, Di Francesco V, Beasley EM;
XX WPI; 2002-608516/65.
XX P-PSDB; AAO18602.
XX
XX New human kinase peptide and nucleic acid molecule, useful for treating
XX disorders associated with abnormal expression of kinase protein, e.g.
XX retinoblastoma, Wilm's tumor, in drug screening assays and
XX

QY 1921 TACATTGTCTAGTGAATGGGAGCCATGAGTGTGTGGAGACACCGGAGGTGTGACG 1980
 Db 1921 TACATTGTCTAGTGAATGGGAGCCATGAGTGTGTGGAGACACCGGAGGTGTGACG 1980
 QY 1981 GAGCTCAAGGCTCGGGAGAGGGCGGCGCAGCCTGAGTGTGTGCTGTGCTGCCAGC 2040
 Db 1981 GAGCTCAAGGCTCGGGAGAGGGCGGCGCAGCCTGAGTGTGTGCTGTGCTGCCAGC 2040
 QY 2041 TCTAGACTGCCAGCTTGGGGAGCCGCGCCGCTGCTGTGGGCCCCAGGGGGCTTCTA 2100
 Db 2041 TCTAGACTGCCAGCTTGGGGAGCCGCGCCGCTGCTGTGGGCCCCAGGGGGCTTCTA 2100
 QY 2101 AGGAGCCAGGGAGCATGTTTCAAGACCCCGGATCCACGTGGGCGCAGTCCCGGGCC 2160
 Db 2101 AGGAGCCAGGGAGCATGTTTCAAGACCCCGGATCCACGTGGGCGCAGTCCCGGGCC 2160
 QY 2161 CTCCTCACTGGAGCCGAAAGGCCAGCAGGCGCAAGCTGGAGGCTGCCCGCCAGCCCTGT 2220
 Db 2161 CTCCTCACTGGAGCCGAAAGGCCAGCAGGCGCAAGCTGGAGGCTGCCCGCCAGCCCTGT 2220
 QY 2221 GCCCAGTGAAGCAGCTCCGCCCTCATCTTGAAGCACCCAGGCTGGCGGTGAGGGCCA 2280
 Db 2221 GCCCAGTGAAGCAGCTCCGCCCTCATCTTGAAGCACCCAGGCTGGCGGTGAGGGCCA 2280
 QY 2281 GGATCCCTGACGCGCTTCCGCTGAGCCTGCTGAGTGGCAGACCGAGCATGCCCTCC 2340
 Db 2281 GGATCCCTGACGCGCTTCCGCTGAGCCTGCTGAGTGGCAGACCGAGCATGCCCTCC 2340
 QY 2341 CCACCCAGAGGACCTCCGGGCAATGCTGCTCCGCTCATGCTGGAGGCTGGCGGTGAGGGCCA 2400
 Db 2341 CCACCCAGAGGACCTCCGGGCAATGCTGCTCCGCTCATGCTGGAGGCTGGCGGTGAGGGCCA 2400
 QY 2401 CTGCTGCTGCCATTAAGACTGGTCCAGACTGCTGTGAAAAAATAAAAAAAAAAAAAA 2460
 Db 2401 CTGCTGCTGCCATTAAGACTGGTCCAGACTGCTGTGAAAAAATAAAAAAAAAAAAAA 2460
 QY 2461 AAAAAAAAAA 2469
 Db 2461 AAAAAAAAAA 2469

RESULT 2

ABV72511
 ID ABV72511 standard; cDNA; 2837 BP.

XX AC ABV72511;
 XX

DT 29-JAN-2003 (first entry)

DE Nucleotide sequence of a human intracellular signalling molecule.

XX Human; intracellular signalling molecule; INTSIG; Alzheimer's disease;
 KW neurodegenerative disorder; Parkinson's disease; muscular disorder;
 KW myotonic dystrophy; catatonia; endocrine disorder; diabetes;
 KW Grave's disease; cancer; leukemia; cervical cancer; breast cancers;
 KW immunological disorder; scleroderma; systemic lupus erythematosus;
 KW allergy; gastrointestinal disorder; Crohn's disease; renal disorder;
 KW Goodpasture's syndrome; infection; cardiovascular disorder;
 KW atherosclerosis; hepatic disease; cirrhosis; gene; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

FT CDS 122..2134

FT /*tag= a

FT /product= "intracellular signalling molecule"

PN WO200277235-A2.

XX 03-OCT-2002.

XX

PF 19-DEC-2001; 2001WO-US050315.

XX
 PR
 PR
 PR
 PR
 PA (INCYTE GENOMICS INC.)
 XX
 PI Burford N, Ding L, Yue H, Thornton M, Walia NK, Gandhi AR;
 PI Arvizu C, Baughn MR, Swarnakar A, Duggan BM, Lu DM, Thangavelu K;
 PI Warren BA, Tang YT, Khan FA, Yao MG, Emerling BM;
 XX
 DR WPI; 2003-018929/01.
 DR P-PSDB; AB899400.
 XX
 PT INTSIG polypeptides, useful for diagnosing, preventing, and treating
 PT disorders associated with an abnormal expression or activity of INTSIG,
 PT e.g. neuromuscular, immunological, cardiovascular disorders, cancer and
 PT infection.
 XX
 PS Claim 5; Page 119-120; 126pp; English.
 CC
 CC The present sequence encodes a human intracellular signalling molecule.
 CC Such molecules are designated INTSIG. The INTSIG polypeptide or its
 CC fragments, and the INTSIG polynucleotide are useful in diagnosing,
 CC preventing, and treating disorders associated with an abnormal expression
 CC or activity of INTSIG, such as neurodegenerative disorders (e.g.
 CC Parkinson's disease, Alzheimer's disease), muscular disorders (e.g.
 CC myotonic dystrophy, catatonia), endocrine disorders (e.g. diabetes,
 CC Grave's disease), cancers (e.g. leukemia, cervical or breast cancers),
 CC immunological disorders (e.g. scleroderma, systemic lupus erythematosus,
 CC allergies), gastrointestinal disorders (e.g. Crohn's disease), renal
 CC disorders (e.g. Goodpasture's syndrome), infections (e.g. viral,
 CC bacterial, fungal, parasitic, protozoal, helminthic), cardiovascular
 CC disorders (e.g. atherosclerosis), or hepatic diseases (e.g. cirrhosis).
 CC INTSIG or its fragments may also be used in screening for compounds that
 CC specifically bind to and modulate the activity of INTSIG. The
 CC polynucleotides can be used to create humanized animals or transgenic
 CC animals to model human disease
 XX
 SQ Sequence 2837 BP; 473 A; 993 C; 908 G; 463 T; 0 U; 0 Other;

Query Match

Best Local Similarity 85.3%; Score 2105; DB 7; Length 2837;

Matches 2116; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 321 GGGCTGTGACTCCCTGACGCGAGATCCAGTGGCGCCAGCTGCAGAGCCGCGAGGCCAGAT 380
 Db 181 GGGCTGTGACTCCCTGACGCGAGATCCAGTGGCGCCAGCTGCAGAGCCGCGAGGCCAGAT 240
 QY 381 TCACCAGCAGATTGACAAGGAGCTGCAGATGGGCGCGCTGAGAACCTCTACAGAGC 440
 Db 241 TCACCAGCAGATTGACAAGGAGCTGCAGATGGGCGCGCGCTGAGAACCTCTACAGAGC 300
 QY 441 CACCAGCAACACCGGCTGAGAGAGACCGTGCCTCGCTGAGCTAGCTACGTCACCTCAA 500
 Db 301 CACCAGCAACACCGGCTGAGAGAGACCGTGCCTCGCTGAGCTAGCTACGTCACCTCAA 360
 QY 501 CCTGCAGCTGCTGAAGGAGGAGCTGGAGAGCTCAGCGGTGCGTGGACCTGCGCGCGCA 560
 Db 361 CCTGCAGCTGCTGAAGGAGGAGCTGGAGAGCTCAGCGGTGCGTGGACCTGCGCGCGCA 420
 QY 561 TGGAGCGAAGCTGTCACTGTCCCTCATGATCCCTCGCTGAGGAGACCAAGGAGCT 620
 Db 421 TGGAGCGAAGCTGTCACTGTCCCTCATGATCCCTCGCTGAGGAGACCAAGGAGCT 480
 QY 621 GGACTGTCTACACCGCTGAAGGAGCTGATCTCAGTGCATCTTTGGAGAGGAGCGGCTC 680
 Db 481 GGACTGTCTACACCGCTGAAGGAGCTGATCTCAGTGCATCTTTGGAGAGGAGCGGCTC 540
 QY 681 CTACGAGGCGAATAATCAGGGAGCTGGAGGCGCTCGGCGAGGCCCATCGGACCCCGCGC 740
 Db 541 CTACGAGGCGAATAATCAGGGAGCTGGAGGCGCTCGGCGAGGCCCATCGGACCCCGCGC 600
 QY 741 GAATGAGTGGGCGCTGGAGGCTGTCTACAGCCCTATTACAAACAGCGTGTCTTCTTGATGC 800

[illegible]

```
PR 20-JUN-2001; 2001US-0299776P.
PR 20-JUN-2001; 2001US-0300001P.
XX (INCY-) INCYTE GENOMICS INC.
XX Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;
PI Dufour GE, Hillman JL, Yu JY, Tuason O, Yap PE, Amshey SR;
PI Daughtery SC, Dam TC, Liu TF, Nguyen DA, Kleefeld Y, Gerstin EH;
PI Peralta CH, David MH, Lewis SA, Chen AJ, Panzer SR, Harris B;
PI Flores V, Marwaha R, Lo A, Lan RY, Urashka ME;
XX WPI; 2003-129518/12.
DR P-FSDB; ABR41761.
XX
XX Novel human diagnostic and therapeutic polypeptide useful for identifying
PT test compound which specifically binds to a polypeptide encoded by human
PT diagnostic and therapeutic polynucleotide, and to induce antibodies.
XX
XX Claim 2; SEQ ID NO 619; 591pp; English.
XX
XX The invention relates to novel human diagnostic and therapeutic
CC polynucleotides designated dithp (ACC46080-ACC46749) and to their encoded
CC proteins (DITHP; ABR41136-ABR41812). The invention also relates to
CC polynucleotide sequences at least 90% identical to the dithp cDNA
CC sequences of the invention; recombinant vectors, host cells and
CC transgenic organisms comprising a dithp nucleic acid sequence; the
CC recombinant production of DITHP proteins; antibodies specific for DITHP
CC proteins; microarrays comprising dithp nucleic acid sequences; methods of
CC detecting dithp nucleotide and protein sequences; methods of screening
CC for compounds which specifically bind a DITHP protein; and methods of
CC assessing the toxicity of test compounds using a dithp hybridisation
CC probe. Dithp nucleic acid sequences and DITHP proteins may be used in the
CC diagnosis of a wide variety of conditions including cancer and other cell
CC proliferative disorders; autoimmune or inflammatory disorders; bacterial,
CC viral, fungal or parasitic infections; hormonal disorders; metabolic
CC disorders; neurological disorders; gastrointestinal disorders; transport
CC disorders; and connective tissue disorders. They may also be used to
CC screen for modulators of protein activity or gene expression. DITHP
CC proteins can additionally be used in analysis of the proteome of a tissue
CC or cell type and to induce antibodies. The dithp nucleic acids are
CC additionally useful in somatic or germline gene therapy of the disorders
CC mentioned above, as a source of antisense sequences, as a source of
CC probes and primers, in genotyping and identification of individuals, in
CC the generation of transgenic animal models of human disease or knock in
CC humanised animals, in toxicological testing, and in transcript imaging.
CC The present sequence represents a dithp cDNA encoding a DITHP protein
CC which is involved in a biochemical pathway. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Query Match 38.6%; Score 952.8; DB 7; Length 3213;
XX Best Local Similarity 93.2%; Pred. No. 9.8e-163;
XX Matches 1018; Conservative 0; Mismatches 72; Indels 2; Gaps 2;
XX
QY 1346 CCCTCTGCGCGGTCTCCCGAGCGACCGAGGGAGAGCTCCCGACCGCAGCAGCAGGTCTTCC 1405
DB 712 CCCCGTGGCGCGACCCCGACGAAAGTGCTGTGATGAGCCCGCCACACGCTTGGCGTTGC 771
QY 1406 TGCAGCCCCCACCCTCTTAAAGCCCCGAGGCCCTGTGCTGCCGAGGAGCTGGAGGAGC 1465
DB 772 CCACCTCTTCTGCCACGTCCCGAGGGCCCGACGGCCACATGGTGTGACATCCCGATGC 831
QY 1466 GCAGGCGCTTGGCAGGACACCTGAAGCTGCCATCTCTGGGCGAGGAGGGCGTGC 1525
DB 832 CCCGCGTGC-AGGCAAGGACACCTGAAGCGTGCCTCTGGGCGAGGAGGGCGCTGC 890
QY 1526 GGCTGCAGCCCTGTGCGCGCTCTTCCGCGAGGTGGACCTGCTTCCGGCTGTGATCTCCC 1585
DB 891 GGCTGCAGCCCTGTGCGCGCTCTTCCGCGAGGTGGACCTGCTTCCGGCTGTGATCTCCC 950
XX
1586 AGACGCTCAGCGCTCACTGCCCAGATATGCGGAGCTCGACCGTGAGGATGACTTCTGTG 1645
DB 951 AGACGCTCAGCGCTCACTGCCCAGATATGCGGAGCTCGACCGTGAGGATGACTTCTGTG 1010
QY 1646 AGGCTGCGGAGGCCCGGACATCCAGCTTAAGCCACAGAGCCAGAGGCCAGGATGC 1705
DB 1011 AGGCTGCGGAGGCCCGGACATCCAGCTTAAGCCACAGAGCCAGAGGCCAGGATGC 1070
QY 1706 CACGCTGCTGCCAGGGGAAAGGGGCTGACATCTTCCATCGGCTGGGGCCCTCTGTGTGT 1765
DB 1071 CACGCTGCTGCCAGGGGAAAGGGGCTGACATCTTCCATCGGCTGGGGCCCTCTGTGTGT 1130
QY 1766 TCTCAGCCAAGAACCGGTGGCGGCTGGTGGGGCCCGTCCACTGACCCGAGGAGGGCG 1825
DB 1131 TCTCAGCCAAGAACCGGTGGCGGCTGGTGGGGCCCGTCCACTGACCCGAGGAGGGCG 1190
QY 1826 GCTTTGGGCTCAGCTTGGGGGAGAGCTGCGCTGTCTCTCATCGCTGCGCTCATTCAGGGA 1885
DB 1191 GCTTTGGGCTCAGCTTGGGGGAGAGCTGCGCTGTCTCTCATCGCTGCGCTCATTCAGGGA 1250
QY 1886 GCCAGGCCGCGCGCTGCGCTGAAAGAGGGCGACTACATTTGTGTGCTGATGAGGCGAGC 1945
DB 1251 GCCAGGCCGCGCGCTGCGCTGAAAGAGGGCGACTACATTTGTGTGCTGATGAGGCGAGC 1310
QY 1946 CATGCAAGTGTGTGAGACACACCGGAGGTGGTGAAGAGCTGAAAGGCTCGCGGAGAGGGCG 2005
DB 1311 CATGCAAGTGTGTGAGACACACCGGAGGTGGTGAAGAGCTGAAAGGCTCGCGGAGAGGGCG 1370
QY 2006 GCGCCAGCTCAGGTGTGTGCTGCTGCTGCCAGCTCTAGACTGCCAGCTTGGGGGAGC 2065
DB 1371 GCGCCAGCTCAGGTGTGTGCTGCTGCTGCCAGCTCTAGACTGCCAGCTTGGGGGAGC 1430
QY 2066 GCGGCCGCTGCTGCTGCGGCGCCAGGGGCTTCTAAGGAGCCAGAGGAGCATGTTGCA 2125
DB 1431 GCGGCCGCTGCTGCTGCGGCGCCAGGGGCTTCTAAGGAGCCAGAGGAGCATGTTGCA 1490
QY 2126 AGACCCCGGATCCACGCTGGGCGGCTGCTGCGGCGCTTCTCAACTGGAGCCGAAAGGCC 2185
DB 1491 AGACCCCGGATCCACGCTGGGCGGCTGCTGCGGCGCTTCTCAACTGGAGCCGAAAGGCC 1550
QY 2186 AGCAGGGCAGACTGGAGGCTGCCCCCGAGCCCTGTGCCCCAGTGAAGCAGCTCCGCCCT 2245
DB 1551 AGCAGGGCAGACTGGAGGCTGCCCCCGAGCCCTGTGCCCCAGTGAAGCAGCTCCGCCCT 1610
QY 2246 CATCTTTGAAGCACCCAGGGTGGCGCTGAGGGCCAGGATCCTTGACGCCCTCAGCCCTG 2305
DB 1611 CATCTTTGAAGCACCCAGGGTGGCGCTGAGGGCCAGGATCCTTGACGCGCTCAGCCCTG 1669
QY 2306 GCTCAGCTGGCAGCAAGCAGCAGCATGCGCTTCCCGACCCAGAGGAGCTCCGGGCAATG 2365
DB 1670 GCTCAGCTGGCAGCAAGCAGCAGCATGCGCTTCCCGACCCAGAGGAGCTCCGGGCAATG 1729
QY 2366 CCTGTCCGCTCATGCTGGAGGCTGCTGCGGCGACCTGCTGCCCATTAAGCTGCTC 2425
DB 1730 CCTGTCCGCTCATGCTGGAGGCTGCTGCGGCGACCTGCTGCCCATTAAGCTGCTC 1789
QY 2426 AGACCTGTCTGA 2437
DB 1790 AGACCTGTCTGA 1801
XX
RESULT 6
ID AA215546 standard; cDNA; 764 BP.
XX
AC AA215546;
XX
DT 12-OCT-1999 (first entry)
XX
DE Human gene expression product cDNA sequence SEQ ID NO:3015.
XX
KW Human; gene; gene expression product; diagnosis; therapy; probe;
KW detection; mapping; tissue typing; profiling; forensic; cancer;
```

KW genetic analysis; colorectal cancer; breast cancer; lung cancer; ss.
XX Homo sapiens.
OS
PN WO9938972-A2.
PD
PD
XX
XX 05-AUG-1999.
XX
XX 28-JAN-1999; 99WO-US001619.
XX
XX 28-JAN-1998; 98US-0072910P.
PR 24-FEB-1998; 98US-0075954P.
PR 31-MAR-1998; 98US-0080114P.
PR 03-APR-1998; 98US-0080515P.
PR 03-APR-1998; 98US-0080666P.
PR 21-OCT-1998; 98US-0105234P.
PR 28-OCT-1998; 98US-0105877P.
XX
XX (CHIR) CHIRON CORP.
PA (HYSR-) HYSRQ INC.
XX
XX Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J;
PI Reinhard C, Glese K, Randazzo F, Kennedy GC, Pot D, Kassam A;
PI Lamson G, Drmanac R, Crkvenjakov R, Dickson M, Drmanac S, Labat I;
PI Leshkowitz D, Kita D, Garcia V, Jones WL, Stache-Crain B;
XX WPI; 1999-494092/41.
XX
XX Novel human genes and their expression products which are differentially
PT expressed in different cell types.
PT
XX
XX Claim 1; Page 1452; 2479pp; English.
XX
XX The present invention describes a library of human polynucleotides
CC comprising the sequences given in AAZ12532 to AAZ17779. Also described is
CC a method of detecting differentially expressed genes correlated with the
CC cancerous state of a mammalian cell, comprising detecting at least one
CC differentially expressed gene product in a test sample from a cell
CC suspected of being cancerous, where the gene product is encoded by one of
CC the 5248 polynucleotide sequences given in AAZ12532 to AAZ17779. The
CC polynucleotides can be used as a source of primers and probes, which can
CC be used for a variety of purpose, e.g. detection of expression levels,
CC mapping, tissue typing or profiling, forensics, genetic analysis and
CC detection of polymorphisms. Polypeptides encoded by the polynucleotides
CC can be used for raising antibodies for experimental, diagnostic and
CC therapeutic purposes. The polynucleotides may also be used to construct
CC arrays for diagnostics (which may be used to determine function of an
CC encoded protein); and to detect differences in expression levels between
CC two cells (e.g. to identify abnormal or diseased tissue in a human, to
CC identify a genetic predisposition or susceptibility to a disease such as
CC cancer). The polynucleotides of the invention are especially used in the
CC diagnosis, prognosis and management of colorectal cancer, breast cancer,
CC and lung cancer. The polynucleotides can also be used to screen for
CC peptide analogues and antagonists
XX
SQ Sequence 764 BP; 157 A; 224 C; 224 G; 137 T; 0 U; 22 Other;

Query Match 21.2%; Score 524.6; DB 2; Length 764;
Best Local Similarity 92.8%; Pred. No. 1.7e-85;
Matches 601; Conservative 0; Mismatches 39; Indels 9; Gaps 5;

QY 1821 GGGCGGCTTTGGCTCACGCTTCGGGAGACTCGCTGTCTCATCGCTGCCGTCAATTCC 1880
DB 51 GGGCGGCTTTGGCTCACGCTTCGGGAGACTCGCTGTCTCATCGCTGCCGTCAATTCC 110
QY 1881 AGGAGCCAGCCGCGCGCGCTGGCTGAAGAGGGGGAGTACATTGTGTCAGTGAATGG 1940
DB 111 AGGGAGCCAGCGCGCGCGCTGGCTGAAGAGGGGGAGTACATTGTGTCAGTGAATGG 170
QY 1941 GCAGCCATGAGTGTGGAGACACACGCGAGGTGGTGACGGAGCTGAAGCTCGGGAGA 2000
DB 171 GCAGCCATGAGTGTGGAGACACACGCGAGGTGGTGACGGAGCTGAAGCTCGGGAGA 230

QY 2001 GGGCGGCGCCAGCTGCAGTGGTGTGCTGCTGCTGCCAGCTCTAGACTGCCAGCTTGGG 2060
DB 231 GGGCGGCGCCAGCTGCAGTGGTGTGCTGCTGCTGCCAGCTCTAGACTGCCAGCTTGGG 290
QY 2061 GGACCGCGCGCGCTCTCTGCTGGGCCCCCAGGGGGCTTCTAAGGAGCCAGAGGAGCATGG 2120
DB 291 GGACCGCGCGCGCTCTCTGCTGGGCCCCCAGGGGGCTTCTAAGGAGCCAGAGGAGCATGG 350
QY 2121 TTGCAAGACCCCGGCTACACGTGGGCGAGTCCCGGGCCCTCTCAACTGGAGCCGAAA 2180
DB 351 TTGCAAGACCCCGGCTACACGTGGGCGAGTCCCGGGCCCTCTCAACTGGAGCCGAAA 410
QY 2181 GGCCAGCAGGGCAAGACTGGAGGCTGCCCGCCAGCCCTGTGCCCAAGTGAAGCCAGCTCC 2240
DB 411 GGCCAGCAGGGCAAGACTGGAGGCTGCCCGCCAGCCCTGTGCCCAAGTGAAGCCAGCTCC 470
QY 2241 GGCCTCATCTTTGAAGACCCAGGGGTGGCGGTGGAGGGCCAGAGTCCCTGACGCCCTCAG 2300
DB 471 G-CCTCATCTTTGAAGACCCAGGGGTGGCGGTGGAGGGCCAGAGTCCCTGACGCC--TCA 527
QY 2301 CCTGTGCTCCAGCTGGCAGCAGCAGCAGCAGTCCCTCCCGCCAGAGGACCTCCGGG 2360
DB 528 CCTGTGCTCCAACTGGCAGCAGCAGCAGCAGTCCCTCCCGCCAGAGGACCTTCNGG 587
QY 2361 CAATGCC-TGTCCTCCGCTCATGCT--GGAGGCTGCTCGGGCAGCTGCC---TSCCATTT 2414
DB 588 CAATGCCTTGTNCCGCTTATGCTTGAAGCTTGTGCTGGGCGACCTTGCCTTGNCCATTT 647
QY 2415 AAGACTGGTCAGACTGTCTGAAAAAATAAAAAAAAAAAAAAAAAAAAAA 2463
DB 648 AAGACTGGTCANAACTGAAAAAATAAAAAAAAAAAAAAAAAAACTTCGAGAAA 696

RESULT 7
AAS15828
ID AAS15828 standard; cDNA; 2100 BP.
XX
AC AAS15828;
XX
DT 16-JAN-2002 (first entry)
XX
DE Human cDNA encoding prostate specific protein PSL22.
XX
KW Human; prostate specific protein; PSL22; prostate cancer;
KW benign prostatic hyperplasia; acute prostatitis; testicular cancer;
KW cryptorchidism; testicular disorder; proliferative disorder; lymphoma;
KW leukaemia; melanoma; ovarian cancer; breast cancer; pancreatic cancer;
KW liver cancer; lung cancer; cytostatic; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT CDS 1..2061
FT /tag= a
FT /product= "PSL22"
XX
XX WO200172962-A2.
XX
XX 04-OCT-2001.
XX
XX 23-MAR-2001; 2001WO-US009410.
XX
XX 24-MAR-2000; 2000US-0191929P.
XX
XX (SAAT/) SAATCIOGLU F.
XX
XX Saatcioglu F;
XX
XX WPI; 2001-662926/76.
XX
XX P-PSDB; AAU10192.
XX
XX New polynucleotide for the diagnosis, prevention and treatment for
PT prostate and testis disorders, particularly prostate cancer, comprises

PT prostate-specific or testis-specific nucleic acids.

XX Claim 4; Fig 14B; 114pp; English.

XX The invention relates to substantially pure prostate-specific or testis-specific polypeptides and the nucleic acids encoding them. Also included are vectors and host cells expressing the proteins, a transgenic animal expressing the protein, antibodies against the proteins, probes for detecting the nucleic acids, antisense molecules for the nucleic acids and methods of isolating modulators of the proteins. Compounds that modulate the prostate specific or testis specific polypeptide are useful to diagnose, prevent or treat disorders of the testis or prostate particularly prostate cancer, benign prostatic hyperplasia, acute prostatitis, testicular cancer, cryptorchidism, undescended, retractile, ascending or vanished testis. Other proliferative disorders for which the modulators may be used include lymphoma, leukaemia, melanoma, ovarian cancer, breast cancer, pancreatic cancer, liver cancer and lung cancer. The present sequence encodes a prostate specific protein, PSU22

XX Sequence 2100 BP; 516 A; 569 C; 584 G; 431 T; 0 U; 0 Other;

Query Match 20.1%; Score 496.4; DB 5; Length 2100;

Best Local Similarity 56.8%; Pred. No. 2.3e-80;

Matches 977; Conservative 0; Mismatches 731; Indels 12; Gaps 3;

QY 321 GGGCTGTGACTCCCTGACGACGATCCAGTGGCGGCGAGCTGCAGCGCGCCAGAT 380

DB 69 GGGCTGTAAATCCCTTGCACAAACCGCGCGAGTAAATTGCAGATCAAGAGCTGCTT 128

QY 381 TCACGACGAGATTGACAGGAGCTGCAGATGCGGCGGCGTGAACCTTCAGAGC 440

DB 129 GAATCAGCAGATCTTGAAGCGCTGCGATGAGGACCGGAGCGGAACCTTCTGAAAGT 188

QY 441 CACCAGCAACACCGGCTGAGAGACGCTGCCCTGGAGCTGAGCTACGCTCACTCCAA 500

DB 189 GGCCCAAACTCAAGGTGCGGACAAAGTGGCGTGGAGCTGAGCTTCGTCACCTCAGA 248

QY 501 CTTGAGCTGTGTAAGGAGGAGCTGGAGAGCTCAGCGGTGGCGCTGGACCTTGGCGGCA 560

DB 249 CTTGAGATGCTCAAGGAAGAGCTGGAGGGGTGAACATCTCGGTGGCGCTATCAGAA 308

QY 561 TGGAGCGAAGCTGTCACTGTCCCATGATCCCCCTGGGCTGAAGAGACCAAGAGCT 620

DB 309 CACAGAGGAGGATTTAGATTCCCTGATTCCTCTGGCCCTGGAAGGAAACGAAGACGT 368

QY 621 GGACTGGTCTACACCGCTGAAGGAGCTGATCTCAGTGCATTTGAGAGGAGCGGCGCTC 680

DB 369 CGACTTTTCAGTCTCTCAGGATTTTATCTGGAACATTTACAGTGAAGATGGCTATTT 428

QY 681 CTACGAGCAGAAATCAGGAGCTGGAGCGCTCGCGGAGGCAATGCGGACCCCGAGCG 740

DB 429 ATATGAAGATGAATTTGAGATCTTATGGATCTGAGACAAGCTTGTGCGAGCGCTAGCG 488

QY 741 GAATGAGTGGGCTGGAGCTGCTCAGAGCTATTACACAGCTGTGCTTCTGATGC 800

DB 489 GGAAGAGCGGCGGTGGAACCTGCTGATGACATCTTCACTCAGCTGGGCTTTGTGAGAG 548

QY 801 GCGCTTCTCTACCCCTGCGAGGAGCTCGGGCTCTTCTTCCACTGGTACGACTGGCTTAC 860

DB 549 TCGATTCTTCGCGCCACACGCGAGATGGACTCTGTTCACCTGGTATGACTCTCTCAC 608

QY 861 TGGGCTCCCGGCCACGAGCGTGGCCCTTCGAGAGGCGAGCGTTCCTTCAACAT 920

DB 609 TGGGCTCCCGGTGAGCCAGCAGAACCTTCTGTGAGAGGCGAGCTGTCTCTTCAACAC 668

QY 921 CGGTGCTTCCACAGCAGATTTGGGCGCGGAGGACCGCTCTCTGCAACGAGGTTGCCG 980

DB 669 TGGGCGCTCTACACCCAGATTGGGACCCGCTGCGATCGGAGCGAGCGTGGGCTGGA 728

QY 981 CCGCGCTATGAGGCGCTTTCAGAGGCGCGCTGGGCGCTTTCAGCGCTCTGAGGAGAACTT 1040

DB 729 GAGTGCCATAGATGCTTTTCAGAGAGCGCGAGGGGTTTTTAATTACCTTGAAGACATTT 788

QY 1041 TTCCCATGGCGGAGCCCGACGATGAGCGCTGCTCCCTCTCGCAGCTGGAGCGCTCAT 1100

DB 789 TACCCATATCTCCAGTTAGACATGAGCCCTGCGCATGCTCAGCGTCTGTCATAATGAT 848

QY 1101 GATGGCCAGGCGCCAGGAATGTGTTTGTAGGGCCCTCTCACACCTGCTCCATGGCCCC 1160

DB 849 GCTTGCACAAAGCCCAAGAAAGCGTGTGTTGAGAAAATCAGCTTCTCTG-----GGATCG 902

QY 1161 CCAAGACTGCTGGCGCCAGCTGCGCTGCGCGAGAGCGCCCGAGGTGGAGCGAGTA 1220

DB 903 GAATGAATTTCTCATCTGTGTGAAGTGGCTCAGAGGCTGCTAAGGTGGAGAGGTCTTA 962

QY 1221 CAGGCTAGTGCACCGGACCATGCGCCAGCCACCCCTCCACGACTAGTGCCTGTCTCTTG 1280

DB 963 CCAACAGCTACACGCGAGCATGAGCGCGCGGTGAAGAGAACATCCCCCTACTCTCTG 1022

QY 1281 GACTGCCCTGGTGATGTCAAGGCGGAGTACTTCCGCTCCCTGGGCCCACTACACGTGAG 1340

DB 1023 GGGCAGCTTAGCTGCGTGAAGGGCCCACTACGCGGCGCTGGGCCCACTACTTCACTGC 1082

QY 1341 CATGGCCCTCTGCGAGCGCTCCCGAGCGACCGAGGAGAGCTCCCGACGCGAGCAGGT 1400

DB 1083 CATCTCTCTATCGACACCAAGGTGAACGAGGCAAGGCTGCGACCAAGAGAGGTG 1142

QY 1401 CTTCTCTGAGC-----CCCCCACTCTCTTAAGCCCCGAGGCGCTGTGC--TGCCCGCAGGA 1454

DB 1143 CTTGCTCCAGCTCTACGACCACTGCGAGGGGCTGACACCTTTGGCCACACTGAGAA 1202

QY 1455 GCTGGAGAGCGCGAGGAGCTTTGGCAAGGACACCTGAAGCGTGCATCTCGGGGAGGA 1514

DB 1203 TGATCAGAGCGCGGACAGCTGGGGAAGTCCCACTTGGCGAGAGCATGGCTCATCA 1262

QY 1515 GAGGCGCTGCGGCTGACGCGCTGTCGCGGCTCTGCGGAGGTGACCTGCTTGGGCG 1574

DB 1263 GAGTCTGCTGGGAGGCGAGCTCTGCAAGAGCTGCGAGCATTTAGGTGCTACAGAA 1322

QY 1575 TGTGATCTCCAGACGCTGCGAGCGCTCACTGCGCAAGTATGCGAGCTCGACCGTGAGGA 1634

DB 1323 GGTGCTGTGTCGCGCACAGGAACGCTCCCGGCTCAGCTACGCCAGACCCAGGAGGAGGA 1382

QY 1635 TGACTTCTGTGAGGCTGCGAGGCGCGGACATCAGCGCTTAAGACCCACGAGAGCCAGA 1694

DB 1383 TGACCTGCTGAACCTGATCGAGCGCCCGCCAGCTTTGTTGCTAAAACTGAGCAAGAGTTGA 1442

QY 1695 GCGCAGGATGCGACGCTGTCAGGGGAAAGGGGCTGACATCTTCCATCGGCTGGGCG 1754

DB 1443 CATTAATTGCCCCAGTTCTCAAGCTGACAGTCAAGGACTTTCTTCCAGAGAGCTGGGCC 1502

QY 1755 CTTGCTGTGTTCTCAGCCAAAGAACCGGTGGCGGTGGTGGGCGCGCTCCACCTGACCCG 1814

DB 1503 CTTATCTGTGTTTTCGGCTAACAGCGGTGAGCGCTCTCTCGAAGCATCCCGTTCACTGC 1562

QY 1815 AGGAGAGGCGGCTTTGGCTCAGCTTTCGGGGAGACATCGGCTGTCTCTCATCGCTGCCGT 1874

DB 1563 AGAAGAGGGGACTTTGGGGTTTCACTTTGAGAGGAAACGCCCGCTTCAGGTTTCACTTCT 1622

QY 1875 CATTCAGGGAGCCAGGCGCGGCTGCGCTGAAGAGGCGGACTACATTGTTGTCAGT 1934

DB 1623 GGATCTTACTGCTCTGCTCGGTGGCAGGAGCCCGGAGAGGAGATTATATTGCTCTCAT 1682

QY 1935 GAATGGGCGAGCATTCAGAGTGGTGAGACACGCGGAGGTGGTGGAGGCTGAAGGCTGC 1994

DB 1683 TCAGCTGTGATTGTTAAGTGGCTGACGCTGAGTGAGTTATGAAGCTGCTGAGAGGCTT 1742

QY 1995 GGGAGAGCGGCGGCCAGCTGCGAGGTGTTGCTGCTGCTG 2034

DB 1743 TGGCGAGGACGAGATCGAGATGAAGTCTGAGCGCTCTCTG 1782

RESULT 8

AAS15827

ID AAS15827 standard; cDNA; 3526 BP.

XX

Db 1323 GGTGCTGTGTCGCGCACAGGAACGCTCCCGGCTACGTCACGCCACCCAGGAGGAGGA 1382
Qy 1635 TGACTTCTGTGAGGTCGCGAGGCGCCCGGACATCAGCCTTAAGACCCACCGAAGCCAGA 1694
Db 1383 TGACCTGCTGAACCTGATCGAGCGCCCGCCAGTGTGTTGCTAAACTGAGCAAGAGGTTGA 1442
Qy 1695 GGCCAGGATGCGCACGCCCTGTCACGAGGGAAGGGGCTGACATCTTCCATCGGCTGGGGCC 1754
Db 1443 CATATATATGCCCCAGTTCTCAAGCTGACAGTCACGGACCTTCTTCCAGAGCTGGGCC 1502
Qy 1755 CTTGCTGTGTTCTCAGCAAGAAACGGTGGCGGCTGGTGGGCGCCGTCACCTGACCGG 1814
Db 1503 CTTATCTGTGTTTTCGGCTTAAACAGCGGTGACGCTCTCTCGAAGCATCCGCTTCACTGC 1562
Qy 1815 AGGAGAGGCGGCTTGGCTTGGCTTGGGAGACTCGCTGCTGCTCATCGCTGCGGT 1874
Db 1563 AGAAGAAGGGGACTTGGGGTTTCACTTGAGAGGGAACGCCCGCTTCAAGTTCACTTCT 1622
Qy 1875 CATTCAGGAGGCGGCGCGCGGCTGCGCTGAAGGAGGCGGCTACATTTGTGTCACT 1934
Db 1623 GGATCCTTACTGCTGCTGCTCGTGGTGGAGAGCGCCGGAAGGAGATTATTTGTCTCCAT 1682
Qy 1935 GAATGGGAGCGATGAGTGTGTGAGACACCGGAGGTGTGACGAGCTGAAGCTGC 1994
Db 1683 TCAGCTTGTGATTGTAGTGGCTGACGCTGAGTGGGTTATGAAGCTGCTGAAGAGCTT 1742
Qy 1995 GGGAGAGCGGCGGCGGCGGCTGCGCTGAGTGGTGTGCTGCTG 2034
Db 1743 TGGCAGGAGCGAGATCGAGATGAAGTGTGAGCTTCCG 1782

RESULT 9

ABX97183
ID ABX97183 standard; cDNA; 2109 BP.

XX AC ABX97183;

XX DT 20-MAY-2003 (first entry)

XX DE Human NOV126b cDNA.

XX KW NOVX; cytostatic; cardiant; antiarteriosclerotic; antiasthmatic; cancer;
hypotensive; cardiomyopathy; bronchial asthma; gene therapy; vaccine;
human; gene; ss.

XX OS Homo sapiens.

XX PN WO200272757-A2.

XX PD 19-SEP-2002.

XX PF 08-MAR-2002; 2002WO-US006908.

XX PR 08-MAR-2001; 2001US-0274101P.

XX PR 08-MAR-2001; 2001US-0274194P.

XX PR 08-MAR-2001; 2001US-0274281P.

XX PR 08-MAR-2001; 2001US-0274322P.

XX PR 09-MAR-2001; 2001US-0274849P.

XX PR 12-MAR-2001; 2001US-0275235P.

XX PR 13-MAR-2001; 2001US-0275578P.

XX PR 13-MAR-2001; 2001US-0275579P.

XX PR 13-MAR-2001; 2001US-0275601P.

XX PR 14-MAR-2001; 2001US-0276000P.

XX PR 16-MAR-2001; 2001US-0276776P.

XX PR 19-MAR-2001; 2001US-0276994P.

XX PR 20-MAR-2001; 2001US-0277239P.

XX PR 20-MAR-2001; 2001US-0277321P.

XX PR 20-MAR-2001; 2001US-0277327P.

XX PR 21-MAR-2001; 2001US-0277791P.

XX PR 22-MAR-2001; 2001US-0277833P.

XX PR 23-MAR-2001; 2001US-0278152P.

XX PR 26-MAR-2001; 2001US-0278894P.

PR 27-MAR-2001; 2001US-0278999P.
PR 27-MAR-2001; 2001US-0279036P.
PR 28-MAR-2001; 2001US-0279344P.
PR 30-MAR-2001; 2001US-0277338P.
PR 30-MAR-2001; 2001US-0279955P.
PR 30-MAR-2001; 2001US-0280233P.
PR 02-APR-2001; 2001US-0280802P.
PR 02-APR-2001; 2001US-0280822P.
PR 02-APR-2001; 2001US-0280900P.
PR 04-APR-2001; 2001US-0281194P.
PR 13-APR-2001; 2001US-0283675P.
PR 30-APR-2001; 2001US-0287424P.
PR 02-MAY-2001; 2001US-0288066P.
PR 03-MAY-2001; 2001US-0288342P.
PR 03-MAY-2001; 2001US-0288528P.
PR 15-MAY-2001; 2001US-0291190P.
PR 16-MAY-2001; 2001US-0291099P.
PR 16-MAY-2001; 2001US-0291240P.
PR 30-MAY-2001; 2001US-0294485P.
PR 31-MAY-2001; 2001US-0294889P.
PR 31-MAY-2001; 2001US-0294899P.
PR 18-JUN-2001; 2001US-0299027P.
PR 19-JUN-2001; 2001US-0299303P.
PR 19-JUN-2001; 2001US-0299310P.
PR 10-JUL-2001; 2001US-0304354P.
PR 31-JUL-2001; 2001US-0309198P.
PR 16-AUG-2001; 2001US-0312903P.
PR 10-SEP-2001; 2001US-0318462P.
PR 12-SEP-2001; 2001US-0318770P.
PR 27-SEP-2001; 2001US-0325430P.
PR 27-SEP-2001; 2001US-0325681P.
PR 18-OCT-2001; 2001US-0330380P.
PR 31-OCT-2001; 2001US-0335301P.
PR 14-NOV-2001; 2001US-0332172P.
PR 14-NOV-2001; 2001US-0332271P.
PR 14-NOV-2001; 2001US-0332272P.
PR 14-NOV-2001; 2001US-0333184P.
PR 14-NOV-2001; 2001US-0333272P.
PR 21-NOV-2001; 2001US-0332094P.
PR 03-DEC-2001; 2001US-0337426P.
PR 03-DEC-2001; 2001US-0338092P.
PR 04-DEC-2001; 2001US-0337185P.
PR 03-JAN-2002; 2002US-0345705P.
PR 07-MAR-2002; 2002US-00092900.

(CURA-) CURAGEN CORP.

Padigaru M, Spytek KA, Shenoy SG, Taupier RJ, Pena CEA, Li L;

Zernhusen BD, Gusev V, Ji W, Gorman L, Miller CE, Kekuda R;

Fatturajan M, Gangolli E, Vernet CAM, Guo X, Tchernev V;

Fernandes ER, Caeman SJ, Malyankar UM, Gerlach V, Liu Y, Anderson D;

Spaderna SK, Catterton E, Burgess C, Leite M, Zhong H, Alsobrook JP;

Lepley DM, Rieger DK;

WPI; 2002-723332/78.

P-PSDB; ABU65216.

NOVX polypeptides and polynucleotides, useful for preventing or treating a disorder associated with aberrant NOVX expression or activity e.g., cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial asthma.

Claim 13; Page 471; 1103pp; English.

This invention describes novel human NOVX polypeptides which have cytostatic, cardiant, antiarteriosclerotic, antiasthmatic and hypotensive activity. Pharmaceutical compositions comprising the NOVX proteins or nucleic acid molecules or NOVX antibodies are useful for preventing or treating a disorder associated with aberrant NOVX expression or activity e.g. cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial asthma. The products of the invention can be used for gene therapy or in a vaccine. ABX97008-ABX97185 are cDNA fragments amplified and isolated by the PCR primers and probes represented in ABX13460-ABX13462 and ABX97186-

CC ABX97593. ABX97008-ABX97185 encode the NOVX proteins described in
CC ABU65041-ABU65218
XX
SQ Sequence 2109 BP; 518 A; 574 C; 585 G; 432 T; 0 U; 0 Other;
Query Match 20.0%; Score 494.8; DB 6; Length 2109;
Best Local Similarity 56.7%; Pred. No. 4.6e-80;
Matches 976; Conservative 0; Mismatches 732; Indels 12; Gaps 3;
QY 321 GGGCTGTCACTCCCTGACGCGATCCAGTCCAGTGGCGCAGCTGCGAGCCGAGGSCCAGAT 380
DB 79 GGGCTGTAAATCCCTTGCACAAACCGCGGAGTAATTTGAGAAATCAAGAGCTGCTTT 138
QY 381 TCACAGCAGATTGACAAAGAGCTGCAGATCGGACGGCGCTGAGAACCTCTACAGAGC 440
DB 139 GAATCAGCAGATCTGAAAGCGTGGATGAGNACCGAGGGGAAACCTTCTGAAGT 198
QY 441 CACAGCMACACCGGCTGAGAGACGGTCCGCTCGAGTGTGAGTACGTTCACTCCAA 500
DB 199 GGCACAACTCAAAGGTGCGGAGCAAGTGGCTGGAGCTGAGCTTCGTCAACTCAGA 258
QY 501 CCTCAGCTGTGAAGAGAGCTGGAGAGCTCAGCGGTGGCGTGGACCTTGGCCGCA 560
DB 259 CTTGAGATGCTCAAGAGAGCTGGAGGGCTGAACATCTCGGTGGCGCTTATCAGAA 318
QY 561 TGGAGCGAAGCTGTCTCACTGTCCCATGATCCCTCGGCTGAGAGACCAAGAGCT 620
DB 319 CACAGAGAGGCAATTTACGATTCCTTCTTCTTGGCTTGAAGNAAAGAGAGCT 378
QY 621 GGACTGTCTACACCGCTGAAGAGCTGATCTCAGTGACATTTGGAGAGGACGGCGCTC 680
DB 379 CGACTTTCAGTCTCTCAAGGATTTTATCCTGGAACATTTACAGTGAAGATGCTATT 438
QY 681 CTACAGCGCAAAATCAGGAGCTGGAGGCTCGGCGAGCCATCGGACCCCGAGCG 740
DB 439 ATATGAAGATGAATTCAGATCTTATGGATCTGAGACAGCTTGTGGAGCGCTAGCCG 498
QY 741 GAATGATTCGGCTGGAGCTGTCTACAGCCATTTACAAACAGCTGTGCTTCTTGGATGC 800
DB 499 GGATGAGCGCGGGTGGAACTGCTGATGACATCTTCACTCAGCTGGCTTGTTCGAGAG 558
QY 801 GCGCTTCTCACCCCTCGAGAGCTCGGGCTCTTCTTCACTGGTACGATCGCTTAC 860
DB 559 TCGAATCTTCCGCGCCACACGCGAGATGGACTCTCTGTTCACCTGGTATGACTCTCTAC 618
QY 861 TGGGCTCCCGCCACGAGCTGCGCTTGGAGAGGCGAGCGCTTCTTCAACAT 920
DB 619 CGGGTTCCGCTCAGCAGCAGAACCTGCTGTGGAGAGCGCAGTGTCTCTTCAACAC 678
QY 921 CGGTGCGCTCCACACGAGATTGGGCGCGCCAGGACCGCTCTGCAACCGAGGCTGCCG 980
DB 679 TGGGGCCCTCTACACCAGATTGGGACCCGCTGCGATCGGACGCGAGCTGGGCTGGA 738
QY 981 CCGCGCTATGGAGCCTTCCAGAGGCGCTGGGCGCTTACGCTCTCAGGAGAACCTT 1040
DB 739 GAGTGCCATAGATGCGCTTTCAGAGAGCCGAGGGGTTTAAATTTACCTGAAAGACAT 798
QY 1041 CTCCTATGCGCCGAGCCAGACATGAGCTGCGTCCCTCTGCGCATGAGGAGAGCTCAT 1100
DB 799 TACCCATCTCAAGTTACGACATGAGCCCTGCCATGCTCAGGTGCTCTCAAAATGAT 858
QY 1101 GATGGCCGAGCCAGGAATGTGTGTTGAGGCGCTCTACACACCTGCCCTCCATGGCCCC 1160
DB 859 GCTTGACAGCCCAAGAAAGCGTGTGAGAAAAATCAGCCTTCTCTG-----GGATCCG 912
QY 1161 CCAAGACTGCTGGCCAGCTGCGCTGGCGAGAGGCGCCCGAGGTGCGACCGAGTA 1220
DB 913 GAATGAATCTTCTATCTGTGAGAGTGGCTCAGGAGGCTGCTAAGTGGGAGAGTCTA 972
QY 1221 CAGGCTAGTGACCGGACCATGCGCCAGCCACCGCTCCAGCATAGTGTGCTGTCTCCTG 1280
DB 973 CCAACAGCTACACGCGCATGAGCCAGGCGCGGTGAAAGAGAACATCCCTTACTCTCTG 1032

QY 1281 GACTGCCCTGTGTGATGTCAAGGCCGAGTACTTTCGCTCCCTTGGGCCCACTACCACTAGC 1340
DB 1033 GGCCAGCTTAGCTGCTGTGAAAGGCCCACTACCGGCGCTTGGGCCCACTACTTCACTGC 1092
QY 1341 CATGGCCCTCTGCGACGGCTCCCAAGCAGCAGGAGAGCTCCCAACGACGACGAGCT 1400
DB 1093 CATCTCTCTCATGACCACTCAGGTGAGCAGGACGAGTCTGGACCACTGAGGAAAGTG 1152
QY 1401 CTTCTGTGAGC-----CCCCACCTCTCTAAAGCCCGAGGCGCTGTGTC--TGCCGACAGA 1454
DB 1153 CTTCTCCAGCTCTACGACCACTATGCGAGAGGGCTGACACCTTTGGCCACACTGAAGAA 1212
QY 1455 GCTGGAGAGCCTCAGGAGCTTGGCAGGACACACTGAGCGTGGCCATCTTGGGCGACGA 1514
DB 1213 TGATCAGAGCGCCGACAGCTGGGGAAGTCCACTTGGCAGAGCCATGGGCTCATCAGA 1272
QY 1515 GGAGCGCTGCGGCTGTCACGCGCTCTGCGCGTCTGCGGAGGTGAGCTTCTTCCGGC 1574
DB 1273 GGAGTCTGCTGCGGAGGCAAGCTCTGCAAGAGCTGCGGAGCATTTGAGGTGCTACAGAA 1332
QY 1575 TGTGATCTCCAGAGCTGCGAGCGCTCACTGCGCAAGTATCGGAGCTTCGACCGTGAGGA 1634
DB 1333 GGTCTGTGTGCGCCACAGGAACTGCTCCCGCTCACGTATCGCCACGACACGAGGAGGA 1392
QY 1635 TGACTTCTGTGAGCTGCGGAGGCGCCGACATCCAGCTTAAGACCCACAGAGCCAGA 1694
DB 1393 TGACTTCTGTGAACTGATTCGACGCGCCCGAGTGTGTTGCTTAAACTTGAGCAAGAGTTGA 1452
QY 1695 GGCCAGGATGCCACCTCTGTCAGGGGAAAGGGCGCTGACATCTTCCATCCGCTGGGCGC 1754
DB 1453 CATATATTGCGCCAGTTCTCCAGCTGACAGTCAAGTCTTCTTCCAGAGCTGGGCGC 1512
QY 1755 CTTCTGTGTCTTACGCAAGAACCGGTGGCGGCTGGTGGGCGCGCTCCACCTGACCGC 1814
DB 1513 CTTATCTGTGTTTTCGGCTTAAAGCGGTGAGCGCTCTCGAAGCATCCGCTTCACTGC 1572
QY 1815 AGGAGAGGCGGCTTTGGCTCTCAGCTTCCGCGGAGACTTCGCTGCTCTCTCATCGCTGCG 1874
DB 1573 AGAAGAGGAGACTTGGGGTTTACCTTGAGAGGAAAGCGCCCGCTTCAAGTTCACTTCT 1632
QY 1875 CATTCAGGAGCAGCGCGCGCTGGCTGAGAGGAGGCGGCTTACATTTGTGTCACT 1934
DB 1633 GGATCTCTTACTGCTCTGCTGCGTGGCAGGAGCCCGGAAAGAGATTATTTGTCTCCAT 1692
QY 1935 GAATGGCAGCCATGAGGTGTGTGGAGACAGCGGAGGTGTGACGAGCTGAAGGCTGC 1994
DB 1693 TCAGCTGTGGAATTTAAGTGGCTGACGCTGAGTGAAGTTATGAAGCTGCTGAAGAGCTT 1752
QY 1995 GGAGAGGCGCGCGCCAGCCTGCAAGTGTGTGCTGCTG 2034
DB 1753 TGGCAGGACGAGATCGAGATGAAGTCTGAGCCTCTG 1792
RESULT 10
ABS71856
ID ABS71856 standard; cdna; 2061 BP.
XX
AC ABS71856;
XX
DT 02-DEC-2002 (first entry)
XX
DE Human GTP-Rho binding protein 2 open reading frame.
XX
KW Human; ss; GTP-Rho binding protein 2; GRBP2; chromosome 19q12; oncogene;
KW tumour; liposarcoma; ichthyosis congenita III;
KW benign familial infantile convulsion; gene therapy.
XX
OS Homo sapiens.
XX
PN EP1231216-A2.
XX
PD 14-AUG-2002.
XX


```
QY 1575 TGTGATCTCCAGACCTGACGGCTCACTGGCCAAAGTATGGAGCTGACCGTGAGGA 1634
Db 1323 GGTGCTGTGCGGCACAGAAAGCTCCCGGCTCAGCTAGCGCCACGACCCAGAGAGGA 1382
QY 1635 TGACTTCTGTAGGCTGCGAGGCCCGGACATCCAGAGCTTAAGACCCACAGAAAGCCAGA 1694
Db 1383 TGACCTGTGAACCTGATGACGCCCCAGTGTGTTGTTAAACTGAGCAAGAGTTGA 1442
QY 1695 GGCACAGATGCCACGGCTGTCACAGGGAGAGGGCCCTGACATCTTCATCGGCTGGGGCC 1754
Db 1443 CATATATATGCCCCAGTTCTCCAAAGCTGACAGTCAAGGACTTCTTCAGAAAGCTGGGCC 1502
QY 1755 CTTGCTGTGTTCTAGCCAGAACCGTGGCGGCTGGTGGGCCCTCCACCTGACCCG 1814
Db 1503 CTTATCTGTGTTTTCGGCTTAACAGCGGTGGAGCCCTCTCGAAGATCGCTTCACTGC 1562
QY 1815 AGAGAGGGCGGCTTTGGGCTCACGCTTCGGGAGACTCGCTGTCTCATCGCTGCGGT 1874
Db 1563 AGAAGAGGGACTTGGGGTTCACTTGAGAGGAACGCCCGTTTCAGGTTCACTTCT 1622
QY 1875 CATTCAGGAGCCAGCGCGCGGCTGCGCTGTAAGAGGGGCGACTACATTTGTGTGAGT 1934
Db 1623 GGATCCTTACTGCTGCTGCTCGGTCGTCAGAGCCCGGAAGGAGATTATATTCTCCAT 1682
QY 1935 GAATGGGCAGCCATCAGGTGGTGAGACACGCGGAGGTGGTGACGAGCTGAAGGCTGC 1994
Db 1683 TCAGCTTGTGATTGTAAGTGGCTGACGCTGAGTGAAGGTATGAAGCTGCTGAAGACTT 1742
QY 1995 GGAGAGGGGGCGCGCCAGGCTGCAGGTGTTGTGCTGCTG 2034
Db 1743 TGSCGAGGACGAGATCGAGATGAAGTCTGTGAGCCTCTG 1782

RESULT 11
ABS71855
ID ABS71855 standard; cDNA; 3484 BP.
XX
AC ABS71855;
XX
DT 02-DEC-2002 (first entry)
XX
DE Human cDNA encoding GTP-Rho binding protein 2.
XX
KW Human; ss; GTP-Rho binding protein 2; GRBP2; chromosome 19q12; oncogene;
KW tumour; liposarcoma; ichthyosis congenita III;
KW benign familial infantile convulsion; gene therapy.
XX
OS Homo sapiens.
XX
PN EF1231216-A2.
XX
PD 14-AUG-2002.
XX
PF 17-JAN-2002; 2002EP-00001026.
XX
PR 30-JAN-2001; 2001WO-US000663.
XX
PR 30-JAN-2001; 2001WO-US000664.
XX
PR 30-JAN-2001; 2001WO-US000665.
XX
PR 30-JAN-2001; 2001WO-US000666.
XX
PR 30-JAN-2001; 2001WO-US000667.
XX
PR 30-JAN-2001; 2001WO-US000668.
XX
PR 30-JAN-2001; 2001WO-US000669.
XX
PR 30-JAN-2001; 2001WO-US000670.
XX
PR 29-JUN-2001; 2001US-00895040.
XX
PA (AEOM-) AEOMICA INC.
XX
PI Shannon ME, J1 Y;
XX
XX WPI; 2002-684026/74.
XX
XX P-PSDB; ABG94709.
XX
```

```
PT Novel GTP-Rho binding protein 2 and nucleic acids encoding the protein,
PT useful for the manufacture of a medicament for treating a disease
PT associated with altered expression or activity of human GRBP2 protein.
XX
PS Claim 1; Fig 3; 101pp; English.
XX
CC The invention relates to an isolated GTP-Rho binding protein 2 (GRBP2)
CC polypeptide or a fragment of at least 6 amino acids or a sequence in
CC which at least 95% of deviations from GRBP2 sequences are conservative
CC substitutions. Also included are an isolated nucleic acid (GRBP2 NA)
CC encoding GRBP2 comprising the full length cDNA or CDS, fragments or
CC variants, GRBP2 vectors, host cells, antibodies, transgenic non-human
CC animals modified to contain GRBP2 NA (or unable to express the endogenous
CC orthologue of GRBP2), diagnosing a disease caused by a mutation in human
CC GRBP2 or altered expression of GRBP2, anti-agonists of GRBP2, GRBP2
CC microarrays, fusion proteins and screening for agents that modulate the
CC expression of GRBP2 NA. GRBP2 is useful for identifying binding partners
CC of GRBP2. GRBP2, GRBP2 NA and Ab are useful in therapy and in the
CC manufacture of a medicament for the treatment or prevention of a disorder
CC associated with increased or decreased expression or activity of human
CC GRBP2 (e.g. tumours, liposarcoma, ichthyosis congenita III and benign
CC familial infantile convulsion, all associated with the chromosome 19
CC location of GRBP2, 19q12). GRBP2 is useful as a standard in immunoassay
CC specific for the proteins, to be used in a therapeutic agent, as
CC vaccines, to be and as antigens (e.g. for epitope mapping) or immunogens
CC (e.g. for raising antibodies. GRBP2 NA is useful as hybridisation probes,
CC to prime synthesis of nucleic acids, to prime first strand cDNA sequence,
CC on an mRNA template, and to drive in vivo expression of the proteins. The
CC vector is useful for shuttling GRBP2 NA between host cells derived from
CC disparate organisms, for inserting GRBP2 NA into host cell chromosome,
CC for expressing sense or antisense RNA transcripts of GRBP2 NA in vitro or
CC within a host cell, and for expressing GRBP2 alone or as fusions to
CC heterologous polypeptides. The antibody is useful as an analytical
CC reagent for detection and quantification of GRBP2 and as an immuno
CC therapeutic agent and is useful for flow cytometric detection, for
CC scanning laser cytometric detection, or for fluorescent immunoassay. The
CC present sequence is a GRBP2 cDNA sequence
XX
SQ Sequence 3484 BP; 932 A; 814 C; 842 G; 896 T; 0 U; 0 Other;
XX
Query Match 20.0%; Score 493.2; DB 6; Length 3484;
Best Local Similarity 56.7%; Pred. No. 9.4e-80;
Matches 975; Conservative 0; Mismatches 733; Indels 12; Gaps 3;
QY 321 GGGCTGTGACTCCTGACGCGAGATCCAGTGGCGCCAGCTGCAGAGCGCGAGGCCCGACAT 380
Db 89 GGGCTGTAAATCCCTTGCACAAACCGCGCGGAGTAATTCGAGAATCAAGAGCTGCTTT 148
QY 381 TCACACGACAGATTGCAAGGAGCTGCAGATGCGGACGGCGCTGAGAACCTCTACAGAGC 440
Db 149 GAATCAGCAGATCCTGAAGCCGTGCGGATGAGATCGGAGCGGAAACCTCTGAAAGT 208
QY 441 CACACAGCAACACCGGCTGAGAGAGACGGTGGCCCTGGAGCTGAGCTAGCTACGTCACTCCAA 500
Db 209 GGCCACAAACTCAAAAGGTGCGGAGCAAGTGGCGGCTGGAGCTGAGCTTCGTCAACTCAGA 268
QY 501 CCTGCAGCTGCTGAAGGAGGAGCTGGAGGAGCTCAGCGGTGGCGTGACCTCGCCGGCA 560
Db 269 CCTGCAGATGCTCAAGGAAGAGCTGGAGGGCTGGAACATCTCGGTGGGGCTCTATCAGAA 328
QY 561 TGGGAGCGAAGCTGTCACTGTCCCCATGATCCCCCTGGGCTGAGAGAGACCAAGAGCT 620
Db 329 CACAGAGGAGGCAATTTACGATTCCCTGTATCTCTTGGCCTGAGGAAACCGAAAGACGT 388
QY 621 GGACTGCTTACACCGCTGAAGGAGCTGATCTCAGTGCATTTGGAGAGACGCGGCCCTC 680
Db 389 CGACTTTGCAGTGTCTCTCAAGGATTTTATCTCGAACATTACAGTGAAGATGGCTATT 448
QY 681 CTACGAGGAGCAAAATCAGGGAGCTGGAGGCCCTTCGGCAGGCCATCGGACCCCGAGCCG 740
Db 449 ATATGAAGATGAATAATGTCAGATCTTATGATCTGAGACAGAGCTTTGTCGAGCGCTAGCCG 508
QY 741 GAATGAGTCGGGCTCGAGCTGCTCAGCCTATTTTCAACAGCTGTGCTTCTGATGC 800
```



```
Db 190 TGCAAGATCAAGAGAGCTGCTTTTGAATCAGCAGATCCTGAAAGCCGTGCGGATGAGGACCG 249
QY 419 GCGCTGAGAACTCTACAGAGCCACCAGCAACAAACCGGGTGAGAGAGACGCTGCCCTGG 478
Db 250 GAGCGGAAAACCTTCTGAAAGTGCGGCACAACTCAAAAGGTGCGGAGCAAGTTCGGCTGG 309
QY 479 AGCTGAGCTACGTCACCTCAACCTCAGCTGCTGAAGGAGGAGCTGGAGGACTCAGCG 538
Db 310 AGCTGAGCTTCGTCACCTCAGCTCAGATGCTCAAGGAGAGCTGGAGGGCTGAACA 369
QY 539 GTGCGCTGGACCCCTGGCGCATGGGAGCGAACTGTCACTGTCCCATGATCCCCCTGG 598
Db 370 TCTCGGTGGCGCTATCAGAAACACAGAGGAGCAATTTACGATTCGCCCTGATTCCTTG 429
QY 599 GCGTGAAGAGACCAAGAGCTGACTGTGCTACACCGCTGAAGGAGCTGATCTCAGTGC 658
Db 430 GCGTGAAGGAAACGAAAGAGCTGCACTTTTGCACTCGTCTCAAGGATTTTATCCTGGAAC 489
QY 659 ACTTTGGAGAGGACGGCGCTCCTACGAGGCAAAATCAGGGAGCTGGAGGGCCCTGCGG 718
Db 490 ATTACAGTGAAGATGCTATTATATGAAGATGAATTCAGATCTTATGATCTGAGAC 549
QY 719 AGCCATGCGGACCCCGAGCGGAATGAGTCGGGCTGGAGCTGCTCAGAGCTATTTACA 778
Db 550 AAGCTTGTGCGAGCGCTAGCGGGATGAGCGCGGGTGGAACTGCTGATGACATCTTCA 609
QY 779 ACCAGCTGTCTTCTGGATGCGGCTTCTCACCCTGCGAGAGGCTCGGGCTCTTCT 838
Db 610 TCCAGCTGGGCTTTGTGAGAGTCGATTTCTTCGCGCCACACCGGCAAGTGGGACTCTGT 669
QY 839 TCCACTGGTACGACTCGCTTACTTGGGTCCCGGCCAGCAGCGCTGCCCTGGCTTCGAGA 898
Db 670 TCACCTGGTATGACTCTCTACCGGGTTCGGGTGAGCAGAGACCTGCTCTGGAGA 729
QY 899 AGGCGACGCTTCTTTCAACATCGTGCCCTCCACAGCAGATTTGGGGCGCCAGAAC 958
Db 730 AGGCCAGTGTCTGTTCAACACTGGGGCCCTTACACCCAGATTTGGGACCGGTGTGATC 789
QY 959 GCTCTCTGCACCGAGGGTGGCGCGCTATGAGGCTTCCAGAGGGCGCTGGGGCCT 1018
Db 790 GGCAGACGCGAGGTGGGTGGAGTGCCATAGTGCCTTTTCAGAGAGCGCGAGGGTTT 849
QY 1019 TCAGCTCTCTGAGGAGAACTTCTCCCATGCGCGCAGCCAGACATGAGCGCTGCGTCCC 1078
Db 850 TAAATTACCTGAAGACACATTTACCCATCTCCAGTTACGACATGAGCCCTGCCATGC 909
QY 1079 TCTGGGACTTGGAGCAGCTCATGATGCCCGAGGCCAGGAATGTGTTTGAAGGCGCTCT 1138
Db 910 TCAGCGTGTCTGTCAAAATGATGCTTGCAAGGCCCAAGAAAGCGTGTGAGAAAATCA 969
QY 1139 CACCACTGCTCCATGGCCCCCAGACTGCTGCGCCAGCTGCGCTCGCGCAGGAGG 1198
Db 970 GCCTTCTG-----GATCCNGAATGAATCTTTCATGCTGGTGAAGTGGCTCAGGAGG 1023
QY 1199 CCGCCAGGTGGCAGCGAGTACAGGCTAGTGACCGGACCATGCGGCCAGCCACCCGCTCC 1258
Db 1024 CTGCTAAGTGGGAGAGTCTACCAACAGCTACACGCGACCATGAGCCAGCGCGGTGA 1083
QY 1259 ACAGCTACGTGCTGCTCTCTGAGCTGCCCTGGTGATGTCAAGCGCGATGACTTCGGCT 1318
Db 1084 AAGAGAACTACCCCTACTCTCTGGGCCAGCTTAGCTGCGTGAAGGCCCAACCATACGCGG 1143
QY 1319 CCTGGGCCACTACACGTAGCATGGCCCTCTGCGACCGCTGCCAGCGCTCCCGACGAGGAG 1378
Db 1144 CCTTGGCCCACTACTTCACTGGCCATCTCTCTCATCGACCAACGAGGTGAAGCCAGGACGG 1203
QY 1379 AGTTCGCCACGACGAGAGGTCTTCTCTGAGC-----CCGCCACTCTCTTAAGCCCGCA 1434
Db 1204 ATCTGGACCCACGAGGAAAGTGCCTGTCGACGCTTACGACCAATGCCAGAGGGGTGA 1263
QY 1435 GGGCTCTGTC--TGCCCCAGAGCTGGAGAGGCGCAGGAGCTTGGCAAGGCACACCTGA 1492
Db 1264 CACCCCTGGCCACACTGAAGAAATGATCAGCAGCGCGCACAGCTGGGGAAATCCCACTTGC 1323
```

```
QY 1493 AGCGTCCCATCTCGGGGAGGAGGCGCTGCGGCTGCACGCGCTGTGCGCGTCTCTGC 1552
Db 1324 GAGAGCCATGGCTCATCAGAGAGTGGTGGCGGAGCGGAGCCTCTGCAAGAAGCTGC 1383
QY 1553 GCAGGTGAGACCTGCTTCGGGCTGTGATCTCCAGACGCTGCAGCGCTCACTGGCCAAAGT 1612
Db 1384 GGACGATTGAGGTGCTTACAGAGGTGCTGTGTGCGCAGACAGGAACGCTCCCGGCTCACGT 1443
QY 1613 ATGCGGAGCTCGACCGCTGAGGATGACTTCTGTAGGCTGCCAGGCCCCGGACATCCAGC 1672
Db 1444 ACGCCAGCAGCCAGGAGGAGGATGACCTGCTGAACCTGATCGACGCCCCCAGTGTGTG 1503
QY 1673 CTAAGACCCACCAAGAGCCAGAGGCCAGGATGCCAGCTGTCCAGGGGAAGGGGCGCTG 1732
Db 1504 CTAAACTGAGCAGAGGTTGACATTAATTATGCCCCAGTTCTCCAAGCTGACAGTCAACGG 1563
QY 1733 ACATTTCCATCGGCTGGGGCCCCCTCTGTGTCTCAGCAGCAAGAACCGGTGCGGCTGG 1792
Db 1564 ACTTCTTCCAGAGCTGGGCCCTTATCTGTGTTTTCGGCTAACAAAGCGGTGACGCGCTC 1623
QY 1793 TGGGGCCGCTCCACCTGACCCGAGGAGAGGGCGGCTTTGGCCTCAGCTTCGGGGAGACT 1852
Db 1624 CTCGAAAGCATCCGCTTCACTGCAAGAAAGAGGAGCTTTGGGGTTTCACTTGAGAGGGAACG 1683
QY 1853 CGCTCTGCTCATCGCTGCGCTCATTTCCAGGGAGCCAGGCGCGCGGCTGGGCTGAAGG 1912
Db 1684 CCCCCTTCAAGTTCACTTCTCGATCCTTACTGTCTGCTGCTGGCAGGAGCCCGG 1743
QY 1913 AGGCGCACTACATTTGTGTGAGTGAATGGGAGCCATGCAAGTGTGTGGAGACACGCGGAGG 1972
Db 1744 AAGGAGATTATATTGTCTCCATTGCTGATTTGTAAGTGGCTGACGCTGAGTGAGG 1803
QY 1973 TGTGTAGCGAGCTGAAGGCTGCGGAGAGCGCGCGCCAGCTTCGAGGTGGTGTGCTGCTGC 2032
Db 1804 TTATGAAGCTGCTGAAGAGCTTTGGCGAGGACGAGATCGAGATGAAGTCTGTGAGCCTCC 1863
QY 2033 TG 2034
Db 1864 TG 1865
```

RESULT 13
ABX97182
ID ABX97182 standard; cDNA; 2310 BP.
XX
AC ABX97182;
XX
DT 20-MAY-2003 (first entry)
XX
DE Human NOV126a cDNA.
XX
KW NOVX; cytostatic; cardiant; antiarteriosclerotic; antiasthmatic; cancer;
KW hypotensive; cardiomyopathy; bronchial asthma; gene therapy; vaccine;
KW human; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200272757-A2.
XX
PD 19-SEP-2002.
XX
PF 08-MAR-2002; 2002WO-US0006908.
XX
PR 08-MAR-2001; 2001US-0274101P.
PR 08-MAR-2001; 2001US-0274194P.
PR 08-MAR-2001; 2001US-0274281P.
PR 08-MAR-2001; 2001US-0274322P.
PR 09-MAR-2001; 2001US-0274849P.
PR 12-MAR-2001; 2001US-0275235P.
PR 13-MAR-2001; 2001US-0275578P.
PR 13-MAR-2001; 2001US-0275579P.
PR 13-MAR-2001; 2001US-0275601P.

```
PR 14-MAR-2001; 2001US-0276000P.
ER 16-MAR-2001; 2001US-0276776P.
PR 19-MAR-2001; 2001US-0276994P.
PR 20-MAR-2001; 2001US-0277239P.
PR 20-MAR-2001; 2001US-0277321P.
PR 20-MAR-2001; 2001US-0277327P.
PR 21-MAR-2001; 2001US-0277791P.
PR 22-MAR-2001; 2001US-0277833P.
PR 23-MAR-2001; 2001US-0278152P.
PR 26-MAR-2001; 2001US-0278894P.
PR 27-MAR-2001; 2001US-0278999P.
PR 28-MAR-2001; 2001US-0279036P.
PR 28-MAR-2001; 2001US-0279344P.
PR 30-MAR-2001; 2001US-0277338P.
PR 30-MAR-2001; 2001US-0279955P.
PR 30-MAR-2001; 2001US-0280233P.
PR 02-APR-2001; 2001US-0280802P.
PR 02-APR-2001; 2001US-0280822P.
PR 02-APR-2001; 2001US-0280900P.
PR 04-APR-2001; 2001US-0281194P.
PR 13-APR-2001; 2001US-0283675P.
PR 30-APR-2001; 2001US-0287424P.
PR 02-MAY-2001; 2001US-0288066P.
PR 03-MAY-2001; 2001US-0288342P.
PR 13-MAY-2001; 2001US-0288528P.
PR 15-MAY-2001; 2001US-0291190P.
PR 16-MAY-2001; 2001US-0291099P.
PR 30-MAY-2001; 2001US-0291240P.
PR 31-MAY-2001; 2001US-0294485P.
PR 31-MAY-2001; 2001US-0294899P.
PR 31-MAY-2001; 2001US-0294899P.
PR 18-JUN-2001; 2001US-0299027P.
PR 19-JUN-2001; 2001US-0299033P.
PR 19-JUN-2001; 2001US-0299310P.
PR 10-JUL-2001; 2001US-0304354P.
PR 31-JUL-2001; 2001US-0309198P.
PR 16-AUG-2001; 2001US-0312903P.
PR 10-SEP-2001; 2001US-0318462P.
PR 12-SEP-2001; 2001US-0318770P.
PR 27-SEP-2001; 2001US-0325430P.
PR 28-SEP-2001; 2001US-0325681P.
PR 18-OCT-2001; 2001US-0330380P.
PR 31-OCT-2001; 2001US-0335301P.
PR 14-NOV-2001; 2001US-0332172P.
PR 14-NOV-2001; 2001US-0332271P.
PR 14-NOV-2001; 2001US-0332272P.
PR 14-NOV-2001; 2001US-0333184P.
PR 14-NOV-2001; 2001US-0333272P.
PR 21-NOV-2001; 2001US-0332094P.
PR 03-DEC-2001; 2001US-0337426P.
PR 03-DEC-2001; 2001US-0338092P.
PR 04-DEC-2001; 2001US-0337185P.
PR 03-JAN-2002; 2002US-0345705P.
PR 07-MAR-2002; 2002US-00092900.
XX
EA (CURA-) CURAGEN CORP.
XX
PI Padigar M, Spytek KA, Shenoy SG, Taupier RJ, Pena CEA, Li L;
PI Zehrusen BD, Gusev V, Ji W, Gorman L, Miller CE, Kekuda R;
PI Patturajan M, Gangolli E, Vernet CAM, Guo X, Tchernev V;
PI Fernandes ER, Casman SE, Malyankar UM, Gerlach V, Liu Y, Anderson D;
PI Spaderna SK, Catterton EJ, Burgess C, Leite M, Zhong H, Alsobrook JP;
PI Lepley DM, Rieger DK;
XX
DR WPI; 2002-723332/78.
DR P-PSDB; ABU65215.
XX
XX NOVX polypeptides and polynucleotides, useful for preventing or treating
PT a disorder associated with aberrant NOVX expression or activity e.g.,
PT cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial
PT asthma.
XX
PS Claim 13; Page 470; 1103pp; English.

XX This invention describes novel human NOVX polypeptides which have
CC cytostatic, cardiant, antiarteriosclerotic, antiasthmatic and hypotensive
CC activity. Pharmaceutical compositions comprising the NOVX proteins or
CC nucleic acid molecules or NOVX antibodies are useful for preventing or
CC treating a disorder associated with aberrant NOVX expression or activity
CC e.g. cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial
CC asthma. The products of the invention can be used for gene therapy or in
CC a vaccine. ABX97008-ABX97185 are cDNA fragments amplified and isolated by
CC the PCR primers and probes represented in ABX13460-ABX13462 and ABX97186-
CC ABX97593. ABX97008-ABX97185 encode the NOVX proteins described in
CC ABU65041-ABU65218
XX
SQ Sequence 2310 BP; 569 A; 622 C; 628 G; 491 T; 0 U; 0 Other;
Query Match 18.8%; Score 463.6; DB 6; Length 2310;
Best Local Similarity 56.7%; Pred. No. 2e-74;
Matches 977; Conservative 0; Mismatches 729; Indels 16; Gaps 6;
QY 321 GGGCTGTGACTCCCTGACGCGAGATCCAGTGGCGCCAGCTGCAGAGCCGAGGCCGACAGAT 380
DB 114 GGGCTGTAAATCCCTTGACAAACCGCGCGAGTAATTCAGAAATCAAGAGCTGCTTT 173
QY 381 TCACGAGCAGATTGACAAAGGAGCTGCAGATCGGACGCGCGCTGAGAACCTCTCAGAGC 440
DB 174 GAATCAGCAGATCTCTGAAAGCCGTGCGGATGAGGACCGGAGCGGAAACCTTCTGAAAGT 233
QY 441 CACCAGCAACAAACCGGGTGACAGAGACGGTCCCTCGAGCTGAGTACGTCACTCCAA 500
DB 234 GGCACAAACTCAAAAGTTCGGGAGCAAGTCCGGCTGGAGCTGAGCTTCGTCAACTCAGA 293
QY 501 CCTGCAGCTGCTGAAGGAGGAGCTGGAGGAGCTCAGCGGCTGCGGTGGACCCCTGGCGCGCA 560
DB 294 CCTGCAGATGCTCAAGGAAGAGCTGGAGGCGCTGAACATCTCGCTGGCGCTTATCAGAA 353
QY 561 TGGGAGCGAAGCTGTCACTGCCCATGATCCCTCGGCTTGGGCTGAAGAGACAGAGGCT 620
DB 354 CACAGAGGAGGCATTTACGATTCCTCTTTGGCTTGAAGGAACGAAAGAGCT 413
QY 621 GGACTGGTCTACACCGCTGAAGGAGCTGATCTCAGTGCACCTTTGGAGAGGACGCGCTC 680
DB 414 CGACTTTGAGTCTGCTCTCAAGGATTTTATCTGGAACATTACAGTGAAGATGGCTATTT 473
QY 681 CTACGAGCAGAAATCAGGGAGCTGGAGGCGCTCGCGGAGGCCCATCGGAGCCCGCAGCG 740
DB 474 ATATGAAGATGAATTCAGATCTTATGGATCTGAGACAAGCTTGTGCGACCGCTAGCG 533
QY 741 GAATGAGTCGGGCTTGGAGCTGCTCAGAGCTTATACAGAGCTGTCTTCTGGATGC 800
DB 534 GGATGAGCGCGGGTGGAACTGCTGATGACATATCTTCATCAGCTGGGCTTTGTGAGAG 593
QY 801 CGGCTTCTCACCCTGCGAGGAGCTCGGGCTCTTCTTCACCTGGTACGACTGCTCTTAC 860
DB 594 TCGATTCTTCCGCGCCACACGCGAGATGGGACTCTGTTCACCTGGTATGACTCTCTCAC 653
QY 861 TGGGCTCCCGCCAGCAGCGTGCCTGGCCCTTCGAGAGGCGCAGCGTCTCTTCAACAT 920
DB 654 CGGGGTTCCGGTTCAGCCAGCAGAACTGCTGCTGGAGAAGGCCAGTGTCTCTTCAACAC 713
QY 921 CGGTGCGCTTCACAGCAGATTTGGGCGCGCGCAGAGCCGCTCTGACCGAGGTTGCCG 980
DB 714 TGGGCGCTCTACACCCAGATTTGGGACCGCGGTGTGATCGGAGACGCGAGGCTGGG 773
QY 981 CCGCGCTATGAGGCGCTTCACAGAGGCGCTGGGCGCTTCAGCGCTCTGAGGAGAACTT 1040
DB 774 GAGTGCCATAGATGCTTTTCAGAGCGCGCAGGGGTTTAAATTTACCTGAAGACACATT 833
QY 1041 CTCCTGCGCGCGCAGCCAGACATGAGCGCTGCGTCCCTCTGCGCACTGGAGAGCTCAT 1100
DB 834 TACCCATACTCCAAGTTACGACATGAGCCCTGCCATGCTCAGCGTGTCTCGTCAAAATGAT 893
QY 1101 GATGCGCCGAGCCAGGAATGTGTTGAGGCGCTCTCACCACTGCTGCTCCATGGCCCC 1160
```

Db 894 GCTTGACAAAGCCAAAGAACGCTGTTTGAGAAAATCAGCCTTCCTG-----GGATCCG 947
QY 1161 CAAAGACTGCTGGCCAGCTGCGCTTGGCGAGAGAGCGCCAGGTCGACCCAGTA 1220
Db 948 GAATGAATTTTCATGCTGGTGAAGTGGCTCAGAGAGGCTGCTAAGTGGGAGAGTCTA 1007
QY 1221 CAGGCTAGTGACCGGACCATGCCCCAGCCACCCGTCACGACTAGCTGCTGCTCTCCTG 1280
Db 1008 CCACAGCTACACGAGCCATGAGCCAGCGCCGGTGAAGAGAAATCCCTTACTCTCTG 1067
QY 1281 GACTGCCCTGGTGCATGTCAAGCCCGAGTACTTCCGCTCCCTGGCCCACTTACACCTAGC 1340
Db 1068 GGCACGCTTAGCTGCTGAAGCCCACTACGCGGCCCTGGCCCACTTACTTCACTGC 1127
QY 1341 CATGGCCCTTGAGACGGCTCCAGCGACCGAGGAGAGAGCTCCCAAGCAGCAGAGGT 1400
Db 1128 CATCTCTCTCATCGACCAACAGCTGAAGCCAGGACCGGATCTGGACCAACAGGAGAGTG 1187
QY 1401 CTTCTCTGCAGC-----CCCCCACTCTCTTAAGCCCCGAGGCCCTGTGC--TGCGCAGGA 1454
Db 1188 CTTGCTCCAGCTCTACGACCAATGCGAGGGGCTGACACCTTGGCCACACTGAGAA 1247
QY 1455 GCTGAGAGCGCAGGAGCTTGGCAAGGCACACCTGAAGCGTGCCATCCTGGGGCAGGA 1514
Db 1248 TGATCAGCAGCGCGCAGCTGGGGAAGTCCCACTTGGCAGAGCCATGGCTCATCAGGA 1307
QY 1515 GAGGCGCTGCGCTGCAAGCCCTGTGCCCGCTCTCGCGGAGGTGACCTGCTTCGGGC 1574
Db 1308 GGAAGTGGTGGCGGAGGCCAGCCCTCTGCAAGAGCTTGGGAGCAATTGAGTGTCTACAGAA 1367
QY 1575 TGTGATCTCCAGACGCTGCGAGCTCTCACTGGGCCAAGTATGCGGAGCTGCACCGTCAGGA 1634
Db 1368 GTGTGCTGTGCGCGCACAGGAAGCTCTCCGGCTCAGTACGCCCGCAGCACAGGAGAGGA 1427
QY 1635 TGACTTCTGAGGCTGCGAGGCCCG--GGAATCAGCCTAAGACCCACACAGAGGCCA 1692
Db 1428 TGACCTGCTGAACCTGATCGACGCCCGCAGAGTGTGTGCTAAAACTGAGCAAGAGTT 1487
QY 1693 GAGGCCAGGATGCCAGCCTGTCCAGGGGAAGGGGCTTACATCTTCCATCGCTGGGG 1752
Db 1488 GACATTATATTGCCCGAGTTCT--CCAGCTACAGTCAACGAGCTTCTTCCAGAGCT--GGG 1545
QY 1753 CCCTGCTGCTGTCTCAGCAAGAACCGGTGGCGGTGTGGGGCCGCTCCACCTGACC 1812
Db 1546 CCCTTATCTGTCTGTGGCTACACAGCGTGAAGCGCTCTCGAAGCATCCGCTTCACT 1605
QY 1813 CGAGGAGAGGGCGGCTTTGGCCTCAGCCTTCCGGGAGACTCGCTGTGCTCTCATCGCTGCC 1872
Db 1606 GCAGAAAGAGGGGACTTGGGTTTACCTTGAGAGGGAACGCCGCCGCTTCAAGTTCACTTC 1665
QY 1873 GTCAATCCAGGAGCCAGCGCGGCTGAGAGGAGCGGCTGAGAGGAGGCTACATTTGTCTCA 1932
Db 1666 CTGGATCTTACTGTCTGTCTCGGTGGAGGAGCCCGGGAAGGAGATTATATTGTCTCC 1725
QY 1933 GTGAATGGGAGCCATGCAAGTGGTGAGACACCGGAGAGGTGTGAGCGAGCTGCAAGGCT 1992
Db 1726 ATTCACTTGTGATTGATGAGTGTGACCTGAGTGAAGTATGAGCTGTCTGAGAGC 1785
QY 1993 GCGGAGAGCGCGGCGCAGCTGCGAGTGGTGTGCTGCTGCTG 2034
Db 1786 TTTGGCGAGGACGAGATCGAGATGAAGTGAAGTGTGAGGCTCTCTG 1827

RESULT 14

AAC91349

ID AAC91349 standard; cDNA; 2757 BP.

XX

AC AAC91349;

XX

DT 16-MAR-2001 (first entry)

XX

DE Human polynucleotide for diagnostics and therapeutics, SEQ ID NO: 49.

XX

KW Human; diagnostics and therapeutics; dithp; cytostatic;
KW immunosuppressive; nootropic; neuroprotective; antiviral; antiallergic;
KW hepatotropic; antidiabetic; antifungal; antiinflammatory; antiulcer; vulnerary;
KW anticonvulsant; antibacterial; antifungal; antiparasitic; cardiac;
KW cancer; immune disorder; cardiovascular disorder; neurological disease;
KW infection; endocrine disorder; metabolic disorder; ss.

XX Homo sapiens.

OS

XX WO200073509-A2.

PN

XX 07-DEC-2000.

PD

XX

XX 31-MAY-2000; 2000WO-US015404.

XX

XX 01-JUN-1999; 99US-0137161P.

PR

XX 02-JUN-1999; 99US-0137109P.

PR

XX 02-JUN-1999; 99US-0137113P.

PR

XX 02-JUN-1999; 99US-0137114P.

PR

XX 02-JUN-1999; 99US-0137173P.

PR

XX 02-JUN-1999; 99US-0137258P.

PR

XX 02-JUN-1999; 99US-0137259P.

PR

XX 02-JUN-1999; 99US-0137260P.

PR

XX 03-JUN-1999; 99US-0137337P.

PR

XX 03-JUN-1999; 99US-0137396P.

PR

XX 03-JUN-1999; 99US-0137411P.

PR

XX 03-JUN-1999; 99US-0137417P.

PR

XX 04-AUG-1999; 99US-0147377P.

PR

XX 04-AUG-1999; 99US-0147436P.

PR

XX 05-AUG-1999; 99US-0147500P.

PR

XX 05-AUG-1999; 99US-0147520P.

PR

XX 05-AUG-1999; 99US-0147527P.

PR

XX 05-AUG-1999; 99US-0147530P.

PR

XX 05-AUG-1999; 99US-0147536P.

PR

XX 05-AUG-1999; 99US-0147541P.

PR

XX 05-AUG-1999; 99US-0147542P.

PR

XX 05-AUG-1999; 99US-0147547P.

PR

XX 05-AUG-1999; 99US-0147549P.

PR

XX 05-AUG-1999; 99US-0147824P.

XX

(INCY-) INCYTE GENOMICS INC.

XX

XX Hodgson DM, Lincoln SE, Russo FD, Spiro PA, Banville SC;

XX Bratcher SR, Dufour GE, Cohen HJ, Rosen BH, Hillman JL;

XX Jones AL, Yu JY, Greenawalt LB, Panzer SR, Roseberry AM, Wright RJ;

XX Daniels SE;

XX

XX WPI; 2001-016511/02.

XX

XX Fifty two human polynucleotides, referred to as DITHP polynucleotides,

XX useful in the diagnosis and treatment of cancer, immune disorders and

XX neurological diseases.

XX

XX Claim 1; Page 261-262; 263pp; English.

XX

The present sequence is one of fifty-two human polynucleotides for
diagnostics and therapeutics. The polynucleotides are referred to as
DITHP polynucleotides. The DITHP polynucleotides and polypeptides are
useful for diagnosing and treating diseases such as cancer, particularly
breast and ovarian cancer, and other cancers of the adrenal gland, bone,
bone marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
immune disorders such as Addison's disease, allergies, autoimmune
hemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis;
cardiovascular disorders such as myocardial ischaemias; neurological
diseases such as cerebral palsy and epilepsy; infectious diseases such as
viral, bacterial, fungal and parasitic infections; and endocrine
disorders (e.g. disorders of hypothalamus), disorders associated with
hypothyroidism and hyperthyroidism, pancreatic disorders (e.g. diabetes
mellitus) and metabolic disorders. The DITHP polynucleotides are useful
for screening for molecules that bind to or are bound by the encoded
polypeptides. The anti-DITHP antibodies are useful in diagnostic assays

1473	QY	GCTTGGCAAGGCACACCTGAAGCGTCCCATCTCTGGGGCAGGAGCGGCTGCGGCTGCA	1533
1015	Db	GCTGGGGAAGTCCCACTTGGCAGAGCCATGGCTCATCANGAGAGTCTCGTGGGAGGC	1074
1533	QY	CGCCCTGTGCGCGTCTCTGGGGAGGTGACCTGCTTCGGGCTGTGATCTCCAGACGCT	1592
1075	Db	GAGCCCTCTCAAGAACTCCG-GANATTGAGGTGCTACAGAGGTGTGNGTCCG-CACA	1132
1593	QY	GCAGCGCTCACTGGCCAAAGTATGCGGAGCTCCAGCCGTGAGGATGACTTCTGTGAGGCTG	1652
1133	Db	GGAACGCTCCGCTCAGCTACGCTACGCCAGCACGAGGAGGATGACCTGCTGAACCTGAT	1192
1653	QY	CGAGGCCCGCGACATCCAGCTAAAGACCCACCAAGAGCCAGAGCCAGGATGCCACGCTT	1712
1193	Db	CGAGGCCCGCGAGTGTGCTTAAACTGACGAGAGGTTGACATTTATTTGCC-CCATT	1251
1713	QY	GTCCAGGGGGAAGGGCCCTGACATCTTCCATCGCTGGGGCCCTCTGTGTCTTCAGC	1772
1252	Db	CTTCAAGCTGACAGTCAACGACTTCTTCCAGAGCTGGGCCCCCTTATCTGTGTTTCGGC	1311
1773	QY	CAAGAACCGGTGCGGCTGCTGGGGCCGCTCCACCTGACCCGAGGAGAGGGGGCTTTG	1832
1312	Db	TAAACAAGCGTGACGCCCTCTCGAAGCATTCGCTTCACTGCAAGAAGAGGAGCTTGG	1371
1833	QY	CTTACGCTTTCGGGGAGACTCGCTGCTCTCATCGCTGCCGCTCATTTCCAGGGAGCCAGG	1892
1372	Db	GTTCACCTTTGAGAGGGAACGCCCGCTTCAGGTTCACTTCTCTGGATCCTTACTGCTCTG	1431
1893	QY	CGCGCGGCTGGCTGAGAGGGGCGACTACATTTGTGTGAGTGAATGGGAGCCATGCGAG	1952
1432	Db	CTCGTGGGAGGAGCCGGGAAGAGATTTATTTGTCTCCATTTCAGCTTTGTGATTGTAA	1491
1953	QY	GTGCTGAGACACGCGGAGTGTGTGACGAGCTGGAAGGCTGCGGGAGAGG	2002
1492	Db	GTGGCTGAGCTGAGTGTGAGTGTGAGCTGTGAAGAGCTTTGGGAGG	1541
RESULT 15			
ABK92255	ID	ABK92255 standard; DNA; 1671 BP.	
XX	AC	ABK92255;	
XX	DT	15-AUG-2002 (first entry)	
XX	DE	Prostate cancer-associated DNA sequence #141.	
XX	KW	Prostate cancer; prostate tumour tissue; human; mammal; cytostatic;	
XX	KX	gene therapy; gene; ds.	
OS		Mammalia.	
XX	PN	WO200230268-A2.	
XX	PD	18-APR-2002.	
XX	PF	12-OCT-2001; 2001WO-US032045.	
XX	PR	13-OCT-2000; 2000US-00687576.	
PR		08-DEC-2000; 2000US-00733288.	
PR		08-DEC-2000; 2000US-00733742.	
PR		24-JAN-2001; 2001US-0263957P.	
PR		16-MAR-2001; 2001US-0276791P.	
PR		16-MAR-2001; 2001US-0276888P.	
PR		06-APR-2001; 2001US-0281922P.	
PR		24-APR-2001; 2001US-0286214P.	
PR		30-APR-2001; 2001US-00847046.	
PR		04-MAY-2001; 2001US-0288589P.	
XX		(E0SB-) EOS BIOTECHNOLOGY INC.	
PA		Gish KC, Mack DH, Wilson KE, Afari D, Hevezi P;	
PI			

XX WPI: 2002-471335/50.
DR P-PSDB; ABG61936.
XX
PT Detecting a prostate cancer-associated transcript in a cell in a patient,
PT useful for diagnosing prostate cancer (PC) or screening modulators of PC,
PT by determining if prostate cancer-associated genes are expressed in a
PT prostate tissue.
XX
XX Claim 22; Page 417-418; 436pp; English.
XX
CC The present invention relates to methods of detecting a prostate cancer-
CC associated transcript in a cell from a patient. The method comprises
CC contacting a biological sample from the patient with prostate cancer-
CC associated polynucleotides (designated PC genes) that selectively
CC hybridize to a sequence that is at least 80% identical to them. The
CC prostate cancer-associated polynucleotide sequences are differentially
CC expressed in prostate tumour tissue or in prostate cancer and are derived
CC from the tissues of various organisms such as humans or other mammals
CC (e.g. mice, sheep and dogs). The methods of the invention are useful for
CC diagnosing and treating prostate cancer in mammals. The prostate cancer-
CC associated genes are useful for diagnosing or treating prostate cancer,
CC as well as for identifying modulators of prostate cancer or agents that
CC inhibit prostate cancer. The nucleic acid sequences are particularly
CC useful in gene therapy, as a vaccine or in antisense applications.
CC ABK92115-ABK92263 represent prostate cancer-associated polynucleotide
CC sequences
XX
SQ Sequence 1671 BP; 402 A; 459 C; 457 G; 353 T; 0 U; 0 Other;
Query Match 15.4%; Score 379.2; DB 6; Length 1671;
Best Local Similarity 56.0%; Pred. No. 3.3e-59;
Matches 783; Conservative 0; Mismatches 603; Indels 12; Gaps 3;
643 GAGCTGATCTCATGTCATTTGAGAGGACGGCGCTCTTACGAGGAGAAATCAGGAG 702
Db 1 GATTTTATCTGGAACATTACAGTGAAGATGGCTATTATATGAAGATGAATTCAGAT 60
703 CTGAGGCGCTGGGAGGCGCATCGGACCCCGAGCGGAATGATCGGCGCTGGAGCTG 762
Db 61 CTATGATGATCTGAGACAAGCTTGTGAGACGCTTAGCGCGGATGAGCGGGGTGGAATG 130
763 CTCACAGCTTATTAACAACAGCTGTGCTTCTCGATGCGCGCTTCTTCACTCCCTGCGAGG 822
Db 121 CTGATGACATCTTATCATCGAGCTGGGCTTTGTGAGAGTGCATTCTTCCCGCCACACGG 180
823 AGCCTGGGCTCTTCTTCACTGGTAGAGTCTGCTTACTGGGGTCCGGCCCGAGCAGCGT 882
Db 181 CAGATGGGACTCTGTTCACCTGGTATGACTCTCTCACCGGGGTTCGGGTTCAGCCAGCAG 240
883 GCCTCTGGCTTCGAGAAGGCGAGCTTCTTCTTCAACATCGGTGCGCTTCCACACGCGAGATT 942
Db 241 AACCTGTGCTGGAGAAGGCCATGTCTCTTTTCAACACTGGGGCCCTCTACACCCAGATT 300
943 GGGGCGCGCCAGGACCGCTCTCTGACCGAGGGTGC CGCGCGCTATGAGGGCTTCCAG 1002
Db 301 GGGACCGGTGTGATCGGACAGCGAGGCTGGGCTGGAGAGTGCATAGATGCTTTCAG 360
1003 AGGGCGCTGGGCGCTTTCAGCTTCTGAGGAGAACTTCTTCCATGGCGCGGAGCCAGAC 1062
Db 361 AGAGCGCGAGGGGTTTTAAATTAACCTGAAGAAGACACATTTACCCATATCTCCAAGTTAGCAG 420
1063 ATGAGCGCTGCTGCTCTGCGACTGGAGCAGCTCATGATGCGCCAGGCGCCAGGAATGT 1122
Db 421 ATGAGCGCTGCCATGCTCAGCGTGTCTGTCATTAATGATGCTTGCACAGCCCGAGAAAGC 480
1123 GTGTTTGAAGGCGCTCTACCACTGCTCTCATGGCCCCCAAGACTGCTGGCCCCAGCTG 1182
Db 481 GTGTTTGAAGAAATCAGCCCTTCTG-----GGATCGGAATGAATTTCTTCATGCTGGTG 534
1183 CGCTGGCGAGGAGCGCCCGGAGGCTGGAGCGGAGTACAGGCTAGTGACCGGACCATG 1242
Db 535 AAGGTGGCTCAGGAGGCTGCTAAGGTGGGAGAGGTCTACCAACAGCTACACGCGCCATG 594

QY 1243 GCCAGCCACCCGCTCCACGACTAGCTGCTGCTCTCTGGAAGTCCCTGTGTCATGTCAAG 1302
Db 595 AGCCAGCGCGCGGTGAAGAGAAACATCCCCCTACTCTGGGCCAGCTTAGCCTGCGTGAAG 654
QY 1303 GCCAGTACTTTCGCTCCCTGGCCCACTACACAGTAGCATGCGCCCTCTCGACGGCTCC 1362
Db 655 GCCACCACTACGCGGCGCTGGCCCACTACTTCACTGCCATCTCTCTCATCGACCACAG 714
QY 1363 CACGACCGAGGAGAGCTCCCAAGCAGCAGCAGAGGTCTTCTCTGAGC-----CCCCAC 1418
Db 715 GTCAAGCCAGGACCGGATCTGGACCAACAGGAGAGTGCCTGCTCCAGCTCTACGACCA 774
QY 1419 CTCTCTTAAGCCCGAGGCGCTCTGCT--GCCCGAGAGCTGAGGAGCCGAGCAGCTT 1476
Db 775 ATCCAGAGGGGTGACACCTTTGGCCACTGAAGAAATGATCAGACGCGCCACAGCTG 834
QY 1477 GGCAAGGCACACTGAAGCGTGCATCTCTGGGCGAGGAGGCGCTGCGGCTGCGAGCC 1536
Db 835 GGGAGTCCACCTTGCAGAGCCATGGCTCATCACAGGAGTGGTGGGAGGCGCAGC 894
QY 1537 CTGTGCGGCTCTGTGCGGAGGTGACCTGCTTGGGCTGTGATCTCCAGAGCTGCGAG 1596
Db 895 CTCTGCAAGAAGCTGCGGAGCAFTGAGGTGCTACAGAAAGTGTGTGTGCCGACAGGAA 954
QY 1597 CGCTCACTGCGCAAGTATGCGGAGCTCGACCGTGAGGATGACTTCTGTGAGGCTGCCGAG 1656
Db 955 CGCTCCGGCTCAGTACGCCCCAGCACAGGAGAGATGACCTGCTGAACCTGATCGAC 1014
QY 1657 GCCCGGACATCCAGCTTAAGACCCACAGAACCCAGAGCCAGGATGCCACGCTGTCC 1716
Db 1015 GCCCGCAGTGTGTTGCTTAAACTGAGCAAGAGGTTGACATTATATTGCCCAAGTTCTCC 1074
QY 1717 CAGGGAAGGGGCTGACATCTTCATCGGCTGGGCGCTGTCTGTGTCTCAGCCAAAG 1776
Db 1075 AAGCTGACAGTCAAGCACTTCTTCCAGAAAGCTGGGCGCTTATCTGTGTTTTCGGCTAAC 1134
QY 1777 AACCGTGGCGGCTGTGTGGGCGCGCTCCACTCACCGGAGAGGCGGCTTGGGCTC 1836
Db 1135 AAGCGTGGAGCGCTCTCGAAGCATCCGTTTCACTGACAGAAAGGGGACTTGGGGTTC 1194
QY 1837 AGCTTTCGGGAGACTCGCTGTCTCATGCTGCGCTGCATTCAGGGAGCCAGGCGCG 1896
Db 1195 ACCTTGAGAGGGAACGCCCGCTTCAAGTTCACTTCTTGGATCTTACTGCTCTGCTCG 1254
QY 1897 GCGCTGGCTGAGAGGCGGACTACATGTGTGATGATGGGAGCCATGAGTGG 1956
Db 1255 GTGGCAGGAGCCCGGAAAGGAGATTATTTGTCTCCATTGCTGAGTTGTAGTGG 1314
QY 1957 TGGAGACACGCGGAGGTGTGACGAGCTGAAGGCTGCGGAGAGCGGCGGCGCAGCTG 2016
Db 1315 CTGACGCTGAGTGAGGTATGAGCTGCTGAAGAGCTTGGCGGAGGACGAGATCGAGATG 1374
QY 2017 CAGGTGGTGTGCTGTG 2034
Db 1375 AAAGTCGTGAGCCTCTG 1392

Search completed: July 13, 2004, 02:51:03
Job time : 645.213 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 12, 2004, 20:08:35 ; Search time 6334.33 Seconds
(without alignments)
16894.276 Million cell updates/sec

Title: US-10-697-266-1
Perfect score: 2469
Sequence: 1 tcggcgcccggtggtggcg.....aaaaaaaaaaaaaaaaaaaaa 2469

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*

1: gb.ba.*

2: gb.hcg.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sy.*

13: gb.un.*

14: gb.vi.*

15: em.ba.*

16: em.fun.*

17: em.hum.*

18: em.in.*

19: em.mu.*

20: em.om.*

21: em.or.*

22: em.ov.*

23: em.pat.*

24: em.ph.*

25: em.pl.*

26: em.ro.*

27: em.sts.*

28: em.un.*

29: em.vi.*

30: em.htg.hum.*

31: em.htg.inv.*

32: em.htg.other.*

33: em.htg.mus.*

34: em.htg.pln.*

35: em.htg.rod.*

36: em.htg.mam.*

37: em.htg.vrt.*

38: em.sy.*

39: em.htgo.hum.*

40: em.htgo.mus.*

41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2469	100.0	2469	6	AR269115	AR269115 Sequence
2	2469	100.0	2469	6	AX642309	AX642309 Sequence
3	2105	85.3	2837	6	AX675063	AX675063 Sequence
4	2105	85.3	3763	9	BC025767	BC025767 Homo sapi
5	2019	81.8	3647	9	AB067516	AB067516 Homo sapi
6	1869	75.7	2088	9	AY082588	AY082588 Homo sapi
7	1102.2	44.6	1932	10	MMU43194	U33194 Mus musculus
8	1043.8	42.3	2301	10	BC052010	BC052010 Mus musculus
9	952.8	38.6	4729	6	BD183400	BD183400 Novel gen
10	558.8	22.6	3230	4	CFA347749	AJ347749 Canis fam
11	524.6	21.2	764	6	BD214873	BD214873 Novel hum
12	496.4	20.1	3516	9	BC036447	BC036447 Homo sapi
13	494.8	20.0	2061	9	HSA347750	AJ347750 Homo sapi
14	494.8	20.0	3500	6	AX835273	AX835273 Sequence
15	494.8	20.0	3500	9	AK098246	AK098246 Homo sapi
16	487.8	19.8	3019	6	AX077672	AX077672 Sequence
17	485.6	19.7	3397	9	BSM803254	AL831950 Homo sapi
18	477.4	19.3	2233	5	BC044556	BC044556 Danio rer
19	475.8	19.3	2224	5	AY391456	AY391456 Homo rer
20	443.4	18.0	2262	9	AF268032	AF268032 Homo sapi
21	418.2	16.9	2757	6	AX593072	AX593072 Sequence
22	368	14.9	19025	6	AR269116	AR269116 Sequence
23	368	14.9	19025	6	AX642311	AX642311 Sequence
24	368	14.9	179057	2	AF276758	AF276758 Homo sapi
25	368	14.9	179941	9	AC105118	AC105118 Homo sapi
26	351.2	14.2	178343	2	AC024049	AC024049 Homo sapi
27	311	12.6	3798	9	AK126506	AK126506 Homo sapi
28	272.8	11.0	2403	3	AF132025	AF132025 Drosophil
29	168.2	6.8	61188	2	AC135076	AC135076 Homo sapi
30	165.4	6.7	123016	2	AF189001	AF189001 Homo sapi
31	165.4	6.7	124060	9	AC087793	AC087793 Homo sapi
32	165	6.7	165	6	BD213706	BD213706 Novel hum
33	159.8	6.5	179057	2	AF276758	AF276758 Homo sapi
34	149.6	6.1	2300	6	AX833345	AX833345 Sequence
35	149.6	6.1	2300	9	AK095001	AK095001 Homo sapi
36	146.6	5.9	232687	2	AC116393	AC116393 Mus muscu
37	143.4	5.8	261608	2	AC097158	AC097158 Rattus no
38	143.4	5.8	272301	2	AC133673	AC133673 Rattus no
39	127	5.1	3174	2	AC014171	AC014171 Drosophil
40	127	5.1	185200	3	AC092401	AC092401 Drosophil
41	127	5.1	191558	3	AC013431	AC013431 Drosophil
42	127	5.1	303345	3	AE003500	AE003500 Drosophil
43	120.2	4.9	511	9	HSA323305	AJ323305 Homo sapi
44	110	4.5	5234	9	AB025194	AB025194 Homo sapi
45	108.4	4.4	2925	6	AX713572	AX713572 Sequence

ALIGNMENTS

RESULT 1
AR269115
LOCUS AR269115 Sequence 1 from patent US 6500655.
DEFINITION AR269115
ACCESSION AR269115
VERSION AR269115.1 GI:29699963
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2469)
AUTHORS Rusch,D., Ketchum,K.A., Di Francesco,V. and Beasley,E.M.
TITLE Isolated human kinase proteins, nucleic acid molecules encoding human kinase proteins, and uses thereof
JOURNAL Patent: US 6500655-A 1 31-DEC-2002;

2469 bp DNA linear PAT 10-APR-2003

FEATURES		Location/Qualifiers
source		1..2469
		/organism="unknown"
		/mol_type="genomic DNA"
ORIGIN		
Query Match 100.0%; Score 2469; DB 6; Length 2469;		
Best Local Similarity 100.0%; Pred. No. 0;		
Matches 2469; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	1	TCGCGGCCACAGTGGTGGCGGGCGCCCTTAGCCCGGCTCGGAGCGCTCGCGAGCGGCGG 60
DB	1	TCGCGGCCACAGTGGTGGCGGGCGCCCTTAGCCCGGCTCGGAGCGCTCGCGAGCGGCGG 60
QY	61	GCTGGCTAACCCGAGGACCCCCACAGCGAGCGGGTGGCGATGATCCTGGAGAGAGG 120
DB	61	GCTGGCTAACCCGAGGACCCCCACAGCGAGCGGGTGGCGATGATCCTGGAGAGAGG 120
QY	121	CCGACGCGCGGGCGCGGAGGAGACCCCGGGCTGCAGATATCTTAGGAGAAACCC 180
DB	121	CCGACGCGCGGGCGCGGAGGAGACCCCGGGCTGCAGATATCTTAGGAGAAACCC 180
QY	181	AGGAAACACGTGTGAGCTCTTTACGGGGAAGACGGGAAGCCCTGAGACAGTGTGTGG 240
DB	181	AGGAAACACGTGTGAGCTCTTTACGGGGAAGACGGGAAGCCCTGAGACAGTGTGTGG 240
QY	241	TGGAGAGGTTGCGGGTCCACAGAGGGAAGACCCAGTGGCTGTGTCAGCTTGGCCCCATG 300
DB	241	TGGAGAGGTTGCGGGTCCACAGAGGGAAGACCCAGTGGCTGTGTCAGCTTGGCCCCATG 300
QY	301	AATCCGACGCTTCATGCACTGGGCTGTGACTCCCTGACGACGATCCAGTGGCGGCAGCTG 360
DB	301	AATCCGACGCTTCATGCACTGGGCTGTGACTCCCTGACGACGATCCAGTGGCGGCAGCTG 360
QY	361	CAGACCGAGGGCCGAGTTCAACAGCAGATTGACAAAGAGTGCAGATCGGACGGGC 420
DB	361	CAGACCGAGGGCCGAGTTCAACAGCAGATTGACAAAGAGTGCAGATCGGACGGGC 420
QY	421	GCTGAGAACCTTACAGAGCACACAGCAACACCGGTTGAGAGACGGTGCCTCTGGAG 480
DB	421	GCTGAGAACCTTACAGAGCACACAGCAACACCGGTTGAGAGACGGTGCCTCTGGAG 480
QY	481	CTGAGCTACGTAACTCAACCTGCAGCTGCTGTGAAGGAGGAGCTGAGAGCTCAGCGGT 540
DB	481	CTGAGCTACGTAACTCAACCTGCAGCTGCTGTGAAGGAGGAGCTGAGAGCTCAGCGGT 540
QY	541	GGCGTGGACCTTGGCGGCATGGGACGAAGTGTCACTGTCCCATGATCCCTCTGGGC 600
DB	541	GGCGTGGACCTTGGCGGCATGGGACGAAGTGTCACTGTCCCATGATCCCTCTGGGC 600
QY	601	CTGAAGGAGACCAAGGAGCTGGACTGTGTCTACACCGCTGAAGGAGCTGATCTCAGTGCAC 660
DB	601	CTGAAGGAGACCAAGGAGCTGGACTGTGTCTACACCGCTGAAGGAGCTGATCTCAGTGCAC 660
QY	661	TTTGGAGAGGACGGGCGCTCTCAAGAGAGAAATCAGGAGCTGGAGGCGCTGGCGGAG 720
DB	661	TTTGGAGAGGACGGGCGCTCTCAAGAGAGAAATCAGGAGCTGGAGGCGCTGGCGGAG 720
QY	721	GCCATGCGACCCCGAGCGGAATCAGTCCGGGCTGGAGCTGCTCACAGCGCTATTACAAAC 780
DB	721	GCCATGCGACCCCGAGCGGAATCAGTCCGGGCTGGAGCTGCTCACAGCGCTATTACAAAC 780
QY	781	CAGCTGTGTTCTCTGGATGCGGCTTCTTCAACCTCGCAGAGCTCGGGCTCTTCTTC 840
DB	781	CAGCTGTGTTCTCTGGATGCGGCTTCTTCAACCTCGCAGAGCTCGGGCTCTTCTTC 840
QY	841	CACTGCTACGACTCGCTTACTGGGTTCCCGGCCAGACGCTGGCCCTTCAGAGAG 900
DB	841	CACTGCTACGACTCGCTTACTGGGTTCCCGGCCAGACGCTGGCCCTTCAGAGAG 900
QY	901	GGCAGCGTTCTCTTCAACATCGGTCCCTTCCACACGAGATTGGGGCGCGCAGACCGC 960
DB	901	GGCAGCGTTCTCTTCAACATCGGTCCCTTCCACACGAGATTGGGGCGCGCAGACCGC 960
QY	961	TCCTGCAACCGAGGGTGCCTCCCGCGCTTATGGAGGCTTTCAGAGGGCGCTTGGGGCCCTTC 1020
DB	961	TCCTGCAACCGAGGGTGCCTCCCGCGCTTATGGAGGCTTTCAGAGGGCGCTTGGGGCCCTTC 1020
QY	1021	AGCCTCTCTGAGGGAGAACTTCTCCCATGCGCCGAGCCACACATGACGCTGGTGCCTC 1080
DB	1021	AGCCTCTCTGAGGGAGAACTTCTCCCATGCGCCGAGCCACACATGACGCTGGTGCCTC 1080
QY	1081	TGGCCTCTGAGCAGCTCATGATGGCCAGGCCACAGGAATGTGTGTTGAGGGCTCTCA 1140
DB	1081	TGGCCTCTGAGCAGCTCATGATGGCCAGGCCACAGGAATGTGTGTTGAGGGCTCTCA 1140
QY	1141	CCACCTGCTTCATGGGCCCCCAAGACTGCTGGCCAGCTGGCTTGGCGCAGGAGGCC 1200
DB	1141	CCACCTGCTTCATGGGCCCCCAAGACTGCTGGCCAGCTGGCTTGGCGCAGGAGGCC 1200
QY	1201	GCCAGGTGTCAGCCGAGTACAGGCTAGTGACACCGGACCATGCGCCAGCCACCGTCCAC 1260
DB	1201	GCCAGGTGTCAGCCGAGTACAGGCTAGTGACACCGGACCATGCGCCAGCCACCGTCCAC 1260
QY	1261	GACTACGTGCTGTCTCTGAGCTGCCCTGGTGCATGTCAAGGCCGAGTACTTCCCGTCC 1320
DB	1261	GACTACGTGCTGTCTCTGAGCTGCCCTGGTGCATGTCAAGGCCGAGTACTTCCCGTCC 1320
QY	1321	CTGGCCCACTACACGTAAGCTAGTGCACCGGACCATGCGCCAGCCAGCGGAGAG 1380
DB	1321	CTGGCCCACTACACGTAAGCTAGTGCACCGGCTTCCGACGCTTCCGAGCGACCGGAGAG 1380
QY	1381	CTCCCCACGCAACGAGCAGTTCCTGACGCCCCCACTCTCTAAGCCCCGAGGGCCCT 1440
DB	1381	CTCCCCACGCAACGAGCAGTTCCTGACGCCCCCACTCTCTAAGCCCCGAGGGCCCT 1440
QY	1441	GTGCTGCGCAGAGGCTGGAGGAGCGCAGCAGCTTGGCAAGGCACACCTGAAAGCGTGC 1500
DB	1441	GTGCTGCGCAGAGGCTGGAGGAGCGCAGCAGCTTGGCAAGGCACACCTGAAAGCGTGC 1500
QY	1501	ATCTGGGGCAGGAGGAGCGCTGCGGCTGCACGCGCTGTGCGCGCTCTGCGCAGGTG 1560
DB	1501	ATCTGGGGCAGGAGGAGCGCTGCGGCTGCACGCGCTGTGCGCGCTCTGCGCAGGTG 1560
QY	1561	GACCTGCTTGGGCTGTGATCTCCACAGCTGCGAGCGCTCACTGGCCAAAGTATGCGGAG 1620
DB	1561	GACCTGCTTGGGCTGTGATCTCCACAGCTGCGAGCGCTCACTGGCCAAAGTATGCGGAG 1620
QY	1621	CTCACCGTGAAGATGACTTCTGTAGGCTGCGAGGCCCCCGGACATCCAGCCCTAAGACC 1680
DB	1621	CTCACCGTGAAGATGACTTCTGTAGGCTGCGAGGCCCCCGGACATCCAGCCCTAAGACC 1680
QY	1681	CACAGAAAGCCAGAGCCAGGATGCCACGCTGTCCAGGGGAAGGGGCGCTGACATCTTC 1740
DB	1681	CACAGAAAGCCAGAGCCAGGATGCCACGCTGTCCAGGGGAAGGGGCGCTGACATCTTC 1740
QY	1741	CATCGGCTGGGGCCCCCTGTGTGTCTCTCAGCCAAAGAACCGTGGCGGCTGTGGGGGCC 1800
DB	1741	CATCGGCTGGGGCCCCCTGTGTGTCTCTCAGCCAAAGAACCGTGGCGGCTGTGGGGGCC 1800
QY	1801	GTCACCTGACCCGAGGAGGGCGGCTTTGGCTCTACGCTTTCGGGAGACTCGCCTGTGC 1860
DB	1801	GTCACCTGACCCGAGGAGGGCGGCTTTGGCTCTACGCTTTCGGGAGACTCGCCTGTGC 1860
QY	1861	CTCATCGCTCCGCTCATTCAGAGGAGCCAGCGCGCGGCTGGCTTGAAGAGGGCGAC 1920
DB	1861	CTCATCGCTCCGCTCATTCAGAGGAGCCAGCGCGCGGCTGGCTTGAAGAGGGCGAC 1920
QY	1921	TACATTGTGTCAGTGAATGGGACGCCATGAGGTGGTGGAGACACCGGAGGTGGTGAACG 1980
DB	1921	TACATTGTGTCAGTGAATGGGACGCCATGAGGTGGTGGAGACACCGCGGAGGTGGTGAACG 1980
QY	1981	GAGCTGAAGGCTGCGGAGAGGGCGGCGCAGCTGAGGTGGTGGAGTGTGCTGCTGCCACGC 2040
DB	1981	GAGCTGAAGGCTGCGGAGAGGGCGGCGCAGCTGAGGTGGTGGAGTGTGCTGCTGCCACGC 2040

```
QY 2041 TCTAGACTGCCAGCTTGGGGGACCGCCGGCCCGCTCTGCTGGGCCCCAGGGGGCTTCTA 2100
Db 2041 TCTAGACTGCCAGCTTGGGGGACCGCCGGCCCGCTCTGCTGGGCCCCAGGGGGCTTCTA 2100
QY 2101 AGGAGCCAGAGGAGCATGTTTGCAGAGACCCCGGCATCCACGTGGGCCAGTCCCGGGCCC 2160
Db 2101 AGGAGCCAGAGGAGCATGTTTGCAGAGACCCCGGCATCCACGTGGGCCAGTCCCGGGCCC 2160
QY 2161 CTCCTCAACTGAGCGCCGAAGGCCCCAGCAGGGAAGACTGGAGGCTGCCCGCCAGCCCTGT 2220
Db 2161 CTCCTCAACTGAGCGCCGAAGGCCCCAGCAGGGAAGACTGGAGGCTGCCCGCCAGCCCTGT 2220
QY 2221 GCCCAGTGAAGCAGCTCCGGCCCTCATCTTGAAGCAACCCAGAGTGGCCGTGAGGGCCA 2280
Db 2221 GCCCAGTGAAGCAGCTCCGGCCCTCATCTTGAAGCAACCCAGAGTGGCCGTGAGGGCCA 2280
QY 2281 GGATCCCTGCAGCCCTCAGCCCTGGCTCCAGCTGGCAGCAAGCAGCAGCATGCCCTCC 2340
Db 2281 GGATCCCTGCAGCCCTCAGCCCTGGCTCCAGCTGGCAGCAAGCAGCAGCATGCCCTCC 2340
QY 2341 CCACCAGAGGACCTCCGGGCAATGCTGTCCGGCTCATGCTGGAGGCTGCTCGGGCA 2400
Db 2341 CCACCAGAGGACCTCCGGGCAATGCTGTCCGGCTCATGCTGGAGGCTGCTCGGGCA 2400
QY 2401 CCTGCTGCCCATTAAGACTGGTCAGACCTGTCTGAAAAAATAAAAAAAAAAAAAA 2460
Db 2401 CCTGCTGCCCATTAAGACTGGTCAGACCTGTCTGAAAAAATAAAAAAAAAAAAAA 2460
QY 2461 AAAAAAAAAA 2469
Db 2461 AAAAAAAAAA 2469

RESULT 2
AX642309
LOCUS AX642309 2469 bp DNA linear PAT 21-FEB-2003
DEFINITION Sequence 1 from Patent WO02061062.
ACCESSION AX642309
VERSION AX642309.1 GI:28474733
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Rusch,D., Ketchum,K., di Francesco,V. and Beasley,E.M.
TITLE Human protein kinase n-like polypeptide and uses thereof
JOURNAL Patent: WO 02061062-A 1 08-AUG-2002;
PE Corporation (NY) (US)
FEATURES
    source
        1..2469
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 2469; DB 6; Length 2469;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2469; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCGCGGCCAGTGTGTGGGGGGCCCTAGCCCGGCTGGGAGCCCTGGCGAGCGGGG 60
Db 1 TCGCGGCCAGTGTGTGGGGGGCCCTAGCCCGGCTGGGAGCCCTGGCGAGCGGGG 60
QY 61 GCTGCTGACCCCGAGGGACCCCGAGCGCAGCGGTGGCGCATGATCCTGGAGGAGG 120
Db 61 GCTGCTGACCCCGAGGGACCCCGAGCGCAGCGGTGGCGCATGATCCTGGAGGAGG 120
QY 121 CCGGACGGCGGGCGCGCGGAGGAGAGCCCGCGCTGCAGATATCTAGAGAAAAACC 180
Db 121 CCGGACGGCGGGCGCGCGGAGGAGAGCCCGCGCTGCAGATATCTAGAGAAAAACC 180
QY 181 AGGAAACACGTGTGAGCTCTTTACGGGGAAGACGGGAAGGCTGAGAGACGTGTGGG 240
```

```
Db 181 AGGAAACACGTGTGAGCTCTTTACGGGGAAGACGGGAAGGCTGAGAGACGTGTGTGG 240
QY 241 TGGAGAGAGGTGTCGGTCCACAGAGGGGAAGAACCCAGTGCGTGTGCACGTTGGCCCCATG 300
Db 241 TGGAGAGAGGTGTCGGTCCACAGAGGGGAAGAACCCAGTGCGTGTGCACGTTGGCCCCATG 300
QY 301 AATCCGAGCTTCTATGCAGTGGGCTGTGACTCCCTGACGCGCAGATCCAGTGGCGCCAGCTG 360
Db 301 AATCCGAGCTTCTATGCAGTGGGCTGTGACTCCCTGACGCGCAGATCCAGTGGCGCCAGCTG 360
QY 361 CAGAGCCGAGGGCCAGATTTCACAGCAGATTTGAACAAGAGCTGCAGATGCGGAGCGGC 420
Db 361 CAGAGCCGAGGGCCAGATTTCACAGCAGATTTGAACAAGAGCTGCAGATGCGGAGCGGC 420
QY 421 GCTGAGAACTCTACAGAGCCACAGCAACAAACGGGTGAGAGACGCTGCCCTGGAG 480
Db 421 GCTGAGAACTCTACAGAGCCACAGCAACAAACGGGTGAGAGACGCTGCCCTGGAG 480
QY 481 CTGAGCTACGTCAACTCCAACTGCAGCTGTCTGAAGGAGAGCTGGAGAGCTCAGCGGT 540
Db 481 CTGAGCTACGTCAACTCCAACTGCAGCTGTCTGAAGGAGAGCTGGAGAGCTCAGCGGT 540
QY 541 GGCCTGGACCTTGGCCGCGCATGGGAGCGAGCTGTCACTGCCCATGATGCCCTGGGC 600
Db 541 GGCCTGGACCTTGGCCGCGCATGGGAGCGAGCTGTCACTGCCCATGATGCCCTGGGC 600
QY 601 CTGAAGAGAGCAAGAGAGCTGGACTGGTCTACACCGCTGAAGAGAGCTGATCTCAGTGAC 660
Db 601 CTGAAGAGAGCAAGAGAGCTGGACTGGTCTACACCGCTGAAGAGAGCTGATCTCAGTGAC 660
QY 661 TTTGGAGAGGACGGCGCTCTCTACGAGGACAGAAATCAGGAGAGCTGGAGGCCCTCGGGCAG 720
Db 661 TTTGGAGAGGACGGCGCTCTCTACGAGGACAGAAATCAGGAGAGCTGGAGGCCCTCGGGCAG 720
QY 721 GCCATGCGGACCCCGAGCGGAGTGAAGTGGGGCTGGGGCTGGAGCTGCTCAGAGCTATTACAC 780
Db 721 GCCATGCGGACCCCGAGCGGAGTGAAGTGGGGCTGGGGCTGGAGCTGCTCAGAGCTATTACAC 780
QY 781 CAGCTGTCTCTCTGGATGGCGCTTCTCTACCCCTGCCAGAGAGCTCGGGCTCTCTTC 840
Db 781 CAGCTGTCTCTCTGGATGGCGCTTCTCTACCCCTGCCAGAGAGCTCGGGCTCTCTTC 840
QY 841 CACTGGTACGACTCGCTTACTGGGGTCCCGGCCAGCAGCGTGCCTCGCCCTTCGAGAAG 900
Db 841 CACTGGTACGACTCGCTTACTGGGGTCCCGGCCAGCAGCGTGCCTCGCCCTTCGAGAAG 900
QY 901 GGCAGCTTCTCTTCAACATCGGTGCCCTCCACACGAGATTGGGGCGGCGCAGGACCGC 960
Db 901 GGCAGCTTCTCTTCAACATCGGTGCCCTCCACACGAGATTGGGGCGGCGCAGGACCGC 960
QY 961 TCCTGCACCGAGGGTGGCGCGCTATGAGGAGCTTCCAGAGGGCGCTGGGGCTTC 1020
Db 961 TCCTGCACCGAGGGTGGCGCGCTATGAGGAGCTTCCAGAGGGCGCTGGGGCTTC 1020
QY 1021 AGCTCTCTGAGGAGAACTTCTCCATGGCGGAGCCCGAGACATGAGCGCTGCGTCCCTC 1080
Db 1021 AGCTCTCTGAGGAGAACTTCTCCATGGCGGAGCCCGAGACATGAGCGCTGCGTCCCTC 1080
QY 1081 TGCGCACTGGAGCAGCTCATGATGGCCCGAGCCAGGAATGTGTGTTTGAAGGGCTCTCA 1140
Db 1081 TGCGCACTGGAGCAGCTCATGATGGCCCGAGCCAGGAATGTGTGTTTGAAGGGCTCTCA 1140
QY 1141 CCACCTGCTCATGGGCCCCCAAGACTGCTGGGCCAGCTGCGCTGCGCGCAGGAGGCC 1200
Db 1141 CCACCTGCTCATGGGCCCCCAAGACTGCTGCGCCAGCTGCGCTGCGCGCAGGAGGCC 1200
QY 1201 GCCCAGGTGGCAGCGAGTACAGGCTAGTGCACCGGACCATGGGCCAGCCACCGCTCCAC 1260
Db 1201 GCCCAGGTGGCAGCGAGTACAGGCTAGTGCACCGGACCATGGGCCAGCCACCGCTCCAC 1260
QY 1261 GACTACGTGCCTGTCTCTGAGCTGCCCTGTGTGATGTCAAGGGCGGAGTACTTCGGCTCC 1320
```


QY 801 GGGCTTCTCAGCCCTGCGCAGGAGCTCGGGCTCTTCTTCCACTGGTACGACTCGCTTAC 860
DB |||||
661 GGGCTTCTCAGCCCTGCGCAGGAGCTCGGGCTCTTCTTCCACTGGTACGACTCGCTTAC 720
QY |||||
861 TGGGGTCCCGGCGCCAGCAGCGTGCCTTCCAGAAAGGCGAGCGTCTCTTCAACAT 920
DB |||||
721 TGGGGTCCCGGCGCCAGCAGCGTGCCTTCCAGAAAGGCGAGCGTCTCTTCAACAT 780
QY |||||
921 CGGTGCTCTCCACACGAGATTGGGGCGCGCCAGAGACCGCTCTCTGACCGAGGGTGCCTG 980
DB |||||
781 CGGTGCTCTCCACACGAGATTGGGGCGCGCCAGAGACCGCTCTCTGACCGAGGGTGCCTG 840
QY |||||
981 CGCGCTATGAGAGGCTTCCAGAGGGCGCTGGGGCTTTCAGCTCTCTGAGGGAGAACTT 1040
DB |||||
841 CGCGCTATGAGAGGCTTCCAGAGGGCGCTGGGGCTTTCAGCTCTCTGAGGGAGAACTT 900
QY |||||
1041 CTCCATGCGCCGAGCCAGACATAGCGCTGCGTCCCTCTGCGCACTGGAGCGCTCAT 1100
DB |||||
901 CTCCATGCGCCGAGCCAGACATAGCGCTGCGTCCCTCTGCGCACTGGAGCGCTCAT 960
QY |||||
1101 GATGCGCCAGGCCAGGAATGTGTGTTGAGGGCTCTCACCACTGCTTCCATGGCCCC 1160
DB |||||
961 GATGCGCCAGGCCAGGAATGTGTGTTGAGGGCTCTCACCACTGCTTCCATGGCCCC 1020
QY |||||
1161 CCAAGACTGCTGCGCCAGCTGCGCTGCGCGCAGAGGGCCGCGAGGTGGCGAGTA 1220
DB |||||
1021 CCAAGACTGCTGCGCCAGCTGCGCTGCGCGCAGAGGGCCGCGAGGTGGCGAGTA 1080
QY |||||
1221 CAGGCTAGTGACCGGACCATGCGCCAGCCACCGCTCCACGACTAGTGGCTGTCTCTG 1280
DB |||||
1081 CAGGCTAGTGACCGGACCATGCGCCAGCCACCGCTCCACGACTAGTGGCTGTCTCTG 1140
QY |||||
1281 GACTGCTGTGTGATGCAAGGCGAGTACTTCCGCTTCCGCTGCTGCTGCTGCTGCTG 1340
DB |||||
1141 GACTGCTGTGTGATGCAAGGCGAGTACTTCCGCTTCCGCTGCTGCTGCTGCTGCTG 1200
QY |||||
1341 CATGSCCTCTGCGAGGCTCCCGAGCGACCGAGGAGAGCTCCCGACGACGAGAGT 1400
DB |||||
1201 CATGSCCTCTGCGAGGCTCCCGAGCGACCGAGGAGAGCTCCCGACGACGAGAGT 1260
QY |||||
1401 CTTCTGCGAGCCCGCCACTCTCTTAAGCCCGAGGCGCTGTGCTGCGCAGGAGCTGA 1460
DB |||||
1261 CTTCTGCGAGCCCGCCACTCTCTTAAGCCCGAGGCGCTGTGCTGCGCAGGAGCTGA 1320
QY |||||
1461 GGAGCGAGGAGCTGTGGCAAGGCAACCTGAAGCGTGCCATCTCTGCGGCGAGGAGGCG 1520
DB |||||
1321 GGAGCGAGGAGCTGTGGCAAGGCAACCTGAAGCGTGCCATCTCTGCGGCGAGGAGGCG 1380
QY |||||
1521 GCTGCGGCTGCGAGCGCTGTGCGCGCTCTGCGGAGGTGACCTGCTTGGGCTGTGAT 1580
DB |||||
1391 GCTGCGGCTGCGAGCGCTGTGCGCGCTCTGCGGAGGTGACCTGCTTGGGCTGTGAT 1440
QY |||||
1581 CTCCAGAGCGTGCAGCGCTCACTGGCCAAAGTATGCGGAGCTCGACCTGTAGGATGACTT 1640
DB |||||
1441 CTCCAGAGCGTGCAGCGCTCACTGGCCAAAGTATGCGGAGCTCGACCTGTAGGATGACTT 1500
QY |||||
1641 CTGTGAGGCTGCGAGGCGCGGACATCAGCCTTAAGACCCACCAAGAGCCAGAGGCGAG 1700
DB |||||
1501 CTGTGAGGCTGCGAGGCGCGGACATCAGCCTTAAGACCCACCAAGAGCCAGAGGCGAG 1560
QY |||||
1701 GATGCCAGCGCTGTCAGGAGGAGGGGCTGACATCTTCCATCGGCTGGGCGCCCTGTC 1760
DB |||||
1561 GATGCCAGCGCTGTCAGGAGGAGGGGCTGACATCTTCCATCGGCTGGGCGCCCTGTC 1620
QY |||||
1761 TGTGTTCTCAGCAAGAAACGGTGGCGCTGTGGGCGCGCTGACCTGACCCGAGGAGA 1820
DB |||||
1621 TGTGTTCTCAGCAAGAAACGGTGGCGCTGTGGGCGCGCTGACCTGACCCGAGGAGA 1680
QY |||||
1821 GGGCGGCTTGGCTTCAAGCTTGGGGAGACTCGCTCTCATCGCTGCGCTCATTC 1880
DB |||||
1681 GGGCGGCTTGGCTTCAAGCTTGGGGAGACTCGCTCTCATCGCTGCGCTCATTC 1740
QY |||||
1881 AGGAGGACGAGCGCGGCGGCTGGCGCTGAAGAGGGGCGACTACATGTGTGCTGAATGG 1940

DB |||||
1741 AGGAGGACCGAGCGCGGCGCTGGCTGAAGAGGGCGACTACATTTGTCTCAGTGAATGG 1800
QY |||||
1941 GCAGCCATGCAGCTGTTGAGACACGCGAGGTGTGACGGAGCTGAAGGCTGCGGAGA 2000
DB |||||
1801 GCAGCCATGCAGCTGTTGAGACACGCGAGGTGTGACGGAGCTGAAGGCTGCGGAGA 1860
QY |||||
2001 GCGGGCGCGCAGCTGCGAGGTGTGCTGCTGCGCTGCGCTAGACTGCCAGCTTGGG 2060
DB |||||
1861 GCGGGCGCGCAGCTGCGAGGTGTGCTGCTGCGCTGCGCTAGACTGCCAGCTTGGG 1920
QY |||||
2061 GSAACCGCGCGCGCTGCTGCTGCGGCGCCAGGGGCTTTAAGAGGCGAGGGAGCATGG 2120
DB |||||
1921 GSAACCGCGCGCGCTGCTGCTGCGGCGCCAGGGGCTTTAAGAGGCGAGGGAGCATGG 1980
QY |||||
2121 TTGCAAGACCCCGGCGATCCAGTGGGCGCAGTCCCGGCGCTGTGCCAGTGAAGCCGCGAA 2180
DB |||||
1981 TTGCAAGACCCCGGCGATCCAGTGGGCGCAGTCCCGGCGCTGTCCAAGTGAAGCCGCGAA 2040
QY |||||
2181 GSCCCAGCAGGCGCAAGACTGGAGGCTGCGCCCGAGCGCTGTGCCAGTGAAGCCAGCTCC 2240
DB |||||
2041 GSCCCAGCAGGCGCAAGACTGGAGGCTGCGCCCGAGCGCTGTGCCAGTGAAGCCAGCTCC 2100
QY |||||
2241 GSCCTCATCTTGAAGACCCAGGGTGGCGGTGAGGGCCAGGATCCCTGCAAGCCCTCAG 2300
DB |||||
2101 GSCCTCATCTTGAAGACCCAGGGTGGCGGTGAGGGCCAGGATCCCTGCAAG-CTCAG 2159
QY |||||
2301 CCCTGGCTCAGCTGGCAGCAGCAGCAGGATGCTGCGGCGCTGCTGCGGCGCTGCTGCCATTAAAGAC 2360
DB |||||
2160 CCCTGGCTCAGCTGGCAGCAGCAGCAGGATGCTGCGGCGCTGCTGCGGCGCTGCTGCCATTAAAGAC 2219
QY |||||
2361 CAATGCTGTCCGCTCATGCTGGAGGCTGCTGCGGCGCTGCTGCGGCGCTGCTGCCATTAAAGAC 2420
DB |||||
2220 CAATGCTGTCCGCTCATGCTGGAGGCTGCTGCGGCGCTGCTGCGGCGCTGCTGCCATTAAAGAC 2279
QY |||||
2421 TGGTCAGACCTGTCTGA 2437
DB |||||
2280 TGGTCAGACCTGTCTGA 2296

RESULT 4
BC025767 3763 bp mRNA linear PRI 06-OCT-2003
LOCUS Homo sapiens rhophilin, Rho GTPase binding protein 1, mRNA (cdna
DEFINITION clone MGC:34714 IMAGE:5211354), complete cds.
ACCESSION BC025767
VERSION BC025767.1 GI:19343950
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3763)
AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, P.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Schaefer, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Faney, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzyzinski, M.I., Skalek, U., Smailus, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
2238257
12477932
2 (bases 1 to 3763)
Strausberg, R.
Direct Submission
Submitted (06-MAR-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK
COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@hgrl.nih.gov
Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,
Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,
Turgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>
Series: IRAC Plate: 49 Row: j Column: 5
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 19882236.

FEATURES
source

1..3763
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/Clone="MGC:34714 IMAGE:5211354"
/tissue_types="Blood, adult leukocytes"
/clone_lib="NIH MGC_118"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1..3763
/genes="RHPN1"
/notes="synonyms: KIAA1929, RHPN, ODF5"
/db_xref="LocusID:114822"
102..2114
/codon_start=1
/product="RHPN1 protein"
/protein_id="AAH25767.1"
/db_xref="GI:19343951"
/db_xref="LocusID:114822"
/translations="MILEERPDGAGGEESPRLOQSDLSLTQIOGOLQSRRAQIHOQI
DKELOVTGAENLYRATSNRVRTEVALSLSYNSNLQLLKELELSGVDPRHGS
EAVTVMPILGKETKELDSPLKELISVHFGEDGASYEAELRELEALRQAMRPSR
NESGLELLTYNQLCDLDFLTPARSLGLFFHWYDSLTPAQKRALAFKGSVLIF
NIGALHTQIGARDGRTSGEARRAMEAFQAAAGFSLLRNFNSHAPSDMSAASLCLAL
EQIMMAQCEVFEGISPPAPQDCLAQLRAQEAQAAAEYRLVHRMTAQPPVHD
YVPVSTALVHVHAEYFRSLAHVAMALCGSPATEGELPTHEQVLPQPTSSKPRG
PVLPOELEERROIGKALHRAILGQBEALRLHALCRVLRVDLLRAVISOTLORSIAK
YALDRDDCEAAEAPDIPQKTHQKEPAEMRPLSQKGGPDIPIRLGLPLSVFSKNEP
RLVGPVHTLKEGSGFGLTRGDSFVLIAVIFGSAAGLKEGDIYVSNQPCRWV
RHAEVHTLKAQSGASLQVSLPSRLPSFLLGDRPVLILGPRLLRSOREHCKTTP
ASTWASPRPLNLSRKAQQKGTGGCPQCPAPKPPAPPSLKHFGWP"

gene

CDS

transduction proteins, including Rho-phillin and BRO1"
/db_xref="CDD:pfam03097"
1641..1841
/note="PDZ; Region: Domain present in PSD-95, Dlg, and
ZO-1/2. Also called DHR (Dlg homologous region) or GIGF
(relatively well conserved tetrapeptide in these domains).
Some PDZs have been shown to bind C-terminal polypeptides"
/db_xref="CDD:smart00228"
ORIGIN
Query Match 85.3%; Score 2105; DB 9; Length 3763;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2116; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 321 GGGCTGTGCTCCCTGACGCGAGATCCAGTGCGCCAGCTCAGAGCGCGAGGCCAGAT 380
Db 161 GGGCTGTGATCTCCCTGACGCGAGATCCAGTGCGCCAGCTCAGAGCGCGAGGCCAGAT 220
QY 381 TCACCAGCAGATTGACAAGGAGCTGCAGATGCGGAGCGGCGCTGAGAACCTCTACAGAGC 440
Db 221 TCACCAGCAGATTGACAAGGAGCTGCAGATGCGGAGCGGCGCTGAGAACCTCTACAGAGC 280
QY 441 CACAGCAACACCGGGTGAAGAGAGAGCGTGCCTCGGAGCTGAGCTAGCTACGTCACTCAA 500
Db 281 CACCAGCAACACCGGGTGAAGAGAGAGCGTGCCTCGGAGCTGAGCTAGCTACGTCACTCAA 340
QY 501 CCTGCACTGCTCAAGGAGGAGCTGAGGAGCTCAGCGTGGGTGAGACCTGCGCGGCA 560
Db 341 CCTGCACTGCTCAAGGAGGAGCTGAGGAGCTCAGCGTGGGTGAGACCTGCGCGGCA 400
QY 561 TGGGAGCGAAGCTGTCTCCTCCCATGATCCCTCGGCTGAGGAGACCAAGGAGCT 620
Db 401 TGGGAGCGAAGCTGTCTCCTCCCATGATCCCTCGGCTGAGGAGACCAAGGAGCT 460
QY 621 GACTGTGTTACACCGCTGAAGAGCTGATCTCAGTGCACCTTTGGAGAGAGCGCGCTC 680
Db 461 GACTGTGTTACACCGCTGAAGAGCTGATCTCAGTGCACCTTTGGAGAGAGCGCGCTC 520
QY 681 CTACGAGGAGAAATCAGGAGCTGAGGAGCTTGGAGGCTTGGAGGAGCGGACCCAGCGG 740
Db 521 CTACGAGGAGAAATCAGGAGCTTGGAGGCTTGGAGGAGCGGACCCAGCGG 580
QY 741 GAATGATCGGGCTGAGAGCTGCTCAAGCCTATTACAAACAGCTGTCTCTGTGATGC 800
Db 581 GAATGATCGGGCTGAGAGCTGCTCAAGCCTATTACAAACAGCTGTCTCTGTGATGC 640
QY 801 GCGCTTCTCACCCTGCCAGAGCTTGGGCTTCTTCCACTGGTACGACTCGCTTAC 860
Db 641 GCGCTTCTCACCCTGCCAGAGCTTGGGCTTCTTCCACTGGTACGACTCGCTTAC 700
QY 861 TGGGGTCCCGGCCAGCAGCTGCGCTTGGGCTTTCGAGAAGGAGGAGCGCTTCTTTCAACAT 920
Db 701 TGGGGTCCCGGCCAGCAGCTGCGCTTGGGCTTTCGAGAAGGAGGAGCGCTTCTTTCAACAT 760
QY 921 CGGTGCGCTTCCACACGAGATTGGGGCGCGCAGGACCGCTCTCTGCAACGAGGTGCGCG 980
Db 761 CGGTGCGCTTCCACACGAGATTGGGGCGCGCAGGACCGCTCTCTGCAACGAGGTGCGCG 820
QY 981 CCGGGCTATGAGAGCTTCCAGAGGCGCTGGGGCTTCCAGCTCTCTGAGGAGAACTT 1040
Db 821 CCGGGCTATGAGAGCTTCCAGAGGCGCTGGGGCTTCCAGCTCTCTGAGGAGAACTT 880
QY 1041 CTCCTATGCGCGAGCCAGACATGAGCGCTGGGCTCTCTTGGGCACTGGAGCAGCTCAT 1100
Db 881 CTCCTATGCGCGAGCCAGACATGAGCGCTGGGCTCTCTTGGGCACTGGAGCAGCTCAT 940
QY 1101 GATGGCCAGCGCCAGGAATGTGTGTTGAGGCGCTCTCAGACCTCTCTCATGGCGCC 1160
Db 941 GATGGCCAGCGCCAGGAATGTGTGTTGAGGCGCTCTCAGACCTCTCTCATGGCGCC 1000
QY 1161 CCAAGACTGCTGCGCCAGCTGCGCGCGAGGAGCGCCCGCAGGTGGAGCGCGAGTA 1220
Db 1001 CCAAGACTGCTGCGCCAGCTGCGCGCGAGGAGCGCCCGCAGGTGGAGCGCGAGTA 1060

misc_feature

/note="Hr1; Region: Hr1 repeat"
/db_xref="CDD:pfam02185"

misc_feature

423..881
/note="BRO1; Region: BRO1-like domain. This functionally
uncharacterised domain is found in a number of signal

```
QY 1221 CAGGCTAGTGCACCGGACCATGGCCAGCACCCGCTCCACGACTAGTGCCTGTCTCTG 1280
Db 1061 CAGGCTAGTGCACCGGACCATGGCCAGCACCCGCTCCACGACTAGTGCCTGTCTCTG 1120
QY 1281 GACTGCCCTGTGTCATGTCACGCGGAGTACTTCGCTCCCTGCGCCACTACACGATAGC 1340
Db 1121 GACTGCCCTGTGTCATGTCACGCGGAGTACTTCGCTCCCTGCGCCACTACACGATAGC 1180
QY 1341 CATGCCCTCTCGGACGGCTCCCGACGACCGAGGAGAGCTCCCGACGACGAGCGT 1400
Db 1181 CATGCCCTCTCGGACGGCTCCCGACGACCGAGGAGAGCTCCCGACGACGAGCGT 1240
QY 1401 CTTCTGACGCCCCCCTCTCTTAAGCCCGAGGCCCTGTGCTGCCAGGAGCTGGA 1460
Db 1241 CTTCTGACGCCCCCCTCTCTTAAGCCCGAGGCCCTGTGCTGCCAGGAGCTGGA 1300
QY 1461 GGAGCGCAGGAGCTGGCAAGGCACCTGAAGCTGTCATCTCTGGGCGCAGGAGGC 1520
Db 1301 GGAGCGCAGGAGCTGGCAAGGCACCTGAAGCTGTCATCTCTGGGCGCAGGAGGC 1360
QY 1521 GCTGCGGCTGTCACGCGCTCTGCGCGCTCTGCGCGAGGTGACCTGCTTCGGGCTGTAT 1580
Db 1361 GCTGCGGCTGTCACGCGCTCTGCGCGCTCTGCGCGAGGTGACCTGCTTCGGGCTGTAT 1420
QY 1581 CTTCCAGAGCTGCGAGCGCTCACTGCGCAAGTATGCGGAGCTCGACCGTGAAGTACTT 1640
Db 1421 CTTCCAGAGCTGCGAGCGCTCACTGCGCAAGTATGCGGAGCTCGACCGTGAAGTACTT 1480
QY 1641 CTGTGAGCTGCGGAGGCCCGGACATCCAGCTTAAGACCCACCAAGCCAGAGGCCAG 1700
Db 1481 CTGTGAGCTGCGGAGGCCCGGACATCCAGCTTAAGACCCACCAAGCCAGAGGCCAG 1540
QY 1701 CATGCAAGCTGTCTCCAGGGAAGGGCTGACATCTTCCATCGCTGGGCGCCCTGTC 1760
Db 1541 CATGCAAGCTGTCTCCAGGGAAGGGCTGACATCTTCCATCGCTGGGCGCCCTGTC 1600
QY 1761 TGTGTTCTCAGCAAGACCGGTGCGGTGTGTGGGCGCCGTCACCTGACCCGAGGAGA 1820
Db 1601 TGTGTTCTCAGCAAGACCGGTGCGGTGTGTGGGCGCCGTCACCTGACCCGAGGAGA 1660
QY 1821 GGGCGGCTTGGCTTCACGCTTGGGGGAGCTCGCTGCTCATCGCTGCGCTCATTC 1880
Db 1661 GGGCGGCTTGGCTTCACGCTTGGGGGAGCTCGCTGCTCATCGCTGCGCTCATTC 1720
QY 1881 AGGAGCCAGGCGCGGCGGCTGGCTGAAGGAGGCGGACTACATTGTGTCAAGTAAAG 1940
Db 1721 AGGAGCCAGGCGCGGCGGCTGGCTGAAGGAGGCGGACTACATTGTGTCAAGTAAAG 1780
QY 1941 GCAGCCATGCAAGTGTGAGACACGCGAGGTGTGACGAGCTGAAGGCTGCGGGAGA 2000
Db 1781 GCAGCCATGCAAGTGTGAGACACGCGAGGTGTGACGAGCTGAAGGCTGCGGGAGA 1840
QY 2001 GCGGGCGCCAGCTGCAAGTGTGCTGCTGCGGAGCTGAGACTGCGGAGCTGGG 2060
Db 1841 GCGGGCGCCAGCTGCAAGTGTGCTGCTGCGGAGCTGAGACTGCGGAGCTGGG 1900
QY 2061 GACCGCGCGGCGGCTGCTGCTGGGCGCCAGGGGGCTTCTAAGGAGCCAGAGGAGCATGG 2120
Db 1901 GACCGCGCGGCGGCTGCTGCTGGGCGCCAGGGGGCTTCTAAGGAGCCAGAGGAGCATGG 1960
QY 2121 TTGCAAGACCCCGGATCCAGTGGGCGAGTCCCGGCGCCCTCTCAACTGAGCGCGAAA 2180
Db 1961 TTGCAAGACCCCGGATCCAGTGGGCGAGTCCCGGCGCCCTCTCAACTGAGCGCGAAA 2020
QY 2181 GSCCCAGCAGGCGCAGACTGGAGGCTGCCCCAGCCCTGTGCCCCAGTGAAGCCAGCTCC 2240
Db 2021 GSCCCAGCAGGCGCAGACTGGAGGCTGCCCCAGCCCTGTGCCCCAGTGAAGCCAGCTCC 2080
QY 2241 GCGCTCATCTTGAAGCACCCAGGCTGGCGGTGAGGGCCAGGATCCCTGCAGCGCCCTCAG 2300
Db 2081 GCGCTCATCTTGAAGCACCCAGGCTGGCGGTGAGGGCCAGGATCCCTGCAGC-CCTCAG 2139
```

```
QY 2301 CCTGGCTCCAGTGGCAGCAAGCACCAGCATGCTCCCTCCACCCAGGAGCCTCCGG 2360
Db 2140 CCTGGCTCCAGTGGCAGCAAGCACCAGCATGCTCCCTCCACCCAGGAGCCTCCGG 2199
QY 2361 CAATGCTGTGCTGGCTCATGCTGAGGCTGCTCGGGACCTGCTGCTGCCATTAAGAC 2420
Db 2200 CAATGCTGTGCTGGCTCATGCTGAGGCTGCTCGGGACCTGCTGCTGCCATTAAGAC 2259
QY 2421 TGGTCAGACCTGCTGGA 2437
Db 2260 TGGTCAGACCTGCTGGA 2276
```

```
RESULT 5
AB067516
LOCUS AB067516 3647 bp mRNA linear PRI 15-SEP-2001
DEFINITION Homo sapiens mRNA for KIAA1929 protein, partial cds.
ACCESSION AB067516
VERSION AB067516.1 GI:15620916
KEYWORDS
```

```
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
```

```
REFERENCE 1
AUTHORS Nagase,T., Kikuno,R. and Ohara,O.
TITLE Prediction of the coding sequences of unidentified human genes.
XXI. The complete sequences of 60 new cDNA clones from brain which
code for large proteins
JOURNAL DNA Res. 8 (4), 179-187 (2001)
MEDLINE 21456161
PUBMED 11572484
```

```
REFERENCE 2 (bases 1 to 3647)
AUTHORS Ohara,O., Nagase,T. and Kikuno,R.
TITLE Direct Submission
JOURNAL Submitted (27-JUN-2001) Osamu Ohara, Kazusa DNA Research Institute,
Department of Human Gene Research; 1532-3, Yana, Kisarazu, Chiba
292-0812, Japan (E-mail:cdnaifoo@kazusa.or.jp,
URL: http://www.kazusa.or.jp/huge, Tel:81-438-52-3913,
Fax:81-438-52-3914)
```

```
FEATURES
Location/Qualifiers
1..3647
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="fj08985"
/tissue_type="brain"
/note="vector:pBluescriptII SK plus"
1..3647
/gene="KIAA1929"
/gene="KIAA1929"
/note="Start codon is not identified. fj08985 cDNA clone
for KIAA1929 has 349-bp, 320-bp and 413-bp insertions
after the positions 450, 936 and 1283 of the sequence of
KIAA1929, respectively."
/codon_start=1
/product="KIAA1929 protein"
/protein_id="BAB67822.1"
/db_xref="GI:15620917"
```

```
gene
CDS
1..1233
/note="Start codon is not identified. fj08985 cDNA clone
for KIAA1929 has 349-bp, 320-bp and 413-bp insertions
after the positions 450, 936 and 1283 of the sequence of
KIAA1929, respectively."
/codon_start=1
```

```
ORIGIN
Query Match 81.8%; Score 2019; DB 9; Length 3647;
Best Local Similarity 96.5%; Pred. No. 0;
Matches 2116; Conservative 0; Mismatches 0; Indels 77; Gaps 2;
```


QY	321	GGGCTGTGACTCCCTGACGAGATCCAGTCCGGCCAGCTGCGAGCGCGAGGCCAGAT	380	QY	1364	-----CAGGACCGAGGAGAGCTCC	1384
Db	51	GGGCTGTGACTCCCTGACGAGATCCAGTCCGGCCAGCTGCGAGCGCGAGGCCAGAT	110	Db	1131	GGCCCCCGCGCAGGGCTCACAGCCTCTGTGTCCCCAGCAGCGAGGAGAGCTCC	1190
QY	381	TCACCCAGCATTGACAAGGAGCTGCGAGTGCAGTGCAGCGGGCGCTGAGAACTCTTACAGAGC	440	QY	1385	CCACGACGAGCAGGTTCTTCTCAGCCCCCCTCTCTAAGCCCCCAGGGCCCTGTGC	1444
Db	111	TCACCCAGCATTGACAAGGAGCTGCGAGTGCAGTGCAGCGGGCGCTGAGAACTCTTACAGAGC	170	Db	1191	CCACGACGAGCAGGTTCTTCTCAGCCCCCCTCTCTAAGCCCCCAGGGCCCTGTGC	1250
QY	441	CACACGACAAACCGGGTGAGAGACGGTGCGCCCTGGAGCTGAGCTGACTCAACTCCAA	500	QY	1445	TGCGCAGAGAGCTGAGGAGCGCAGCAGCTTGGCAGGACACACCTGAGGGTGCCTATCC	1504
Db	171	CACACGACAAACCGGGTGAGAGACGGTGCGCCCTGGAGCTGAGCTGACTCAACTCCAA	230	Db	1251	TGCGCAGGAGCTGAGGAGCGCAGCAGCTTGGCAGGACACACCTGAGGGTGCCTATCC	1310
QY	501	CTTGCAGCTGCTGAAAGGAGCTGAGGAGCTCAGCGTGGCTGAGCAGCCTGGCGGCA	560	QY	1505	TGCGGACGAGGAGGCGCTGCGGCTGCAAGCCTGTGCGCGCTCTCTGCGGAGGTGAGC	1564
Db	231	CTTGCAGCTGCTGAAAGGAGCTGAGGAGCTCAGCGTGGCTGAGCAGCCTGGCGGCA	290	Db	1311	TGSGGGCAGGAGGAGGCGCTGCGGCTGCAAGCCTGTGCGCGCTCTCTGCGGAGGTGAGC	1370
QY	561	TGGGAGCGAAGCTGTCACTGTCCCATGATCCCCCTGGGCTGAAGGAGACCAAGAGCT	620	QY	1565	TGCTTCGGGCTGTGATCTCTCCAGACGCTGACGCGCTCACTGGCCAACTATGCGGAGCTCG	1624
Db	291	TGGGAGCGAAGCTGTCACTGTCCCATGATCCCCCTGGGCTGAAGGAGACCAAGAGCT	350	Db	1371	TGCTTCGGGCTGTGATCTCTCCAGACGCTGACGCGCTCACTGGCCAACTATGCGGAGCTCG	1430
QY	621	GGACTGGTCTACACCGCTGAAGAGCTGATCTCAGTGCACTTTGGAGAGACCGCGCTC	680	QY	1625	ACCGTGAGGATGACTTCTGTGAGGCTGCCAGAGCCCGGGACATCCAGCCTTAAGACCCACC	1684
Db	351	GGACTGGTCTACACCGCTGAAGAGCTGATCTCAGTGCACTTTGGAGAGACCGCGCTC	410	Db	1431	ACCGTGAGGATGACTTCTGTGAGGCTGCCAGAGCCCGGGACATCCAGCCTTAAGACCCACC	1490
QY	681	CTAGAGGCAAGAAATCAGGAGCTGGAGGCGCTGCCGAGGCGCAATGCGGACCCCGCCG	740	QY	1685	AGAAGCCAGAGGCGCAGGATGCCACGCTGTCTCCAGAGGAGAGGGGCGCTGACATCTTCCATC	1744
Db	411	CTAGAGGCAAGAAATCAGGAGCTGGAGGCGCTGCCGAGGCGCAATGCGGACCCCGCCG	470	Db	1491	AGAAGCCAGAGGCGCAGGATGCCACGCTGTCTCCAGAGGAGAGGGGCGCTGACATCTTCCATC	1550
QY	741	GAATGAGTGGGCGCTGGAGCTGTCTCAGAGCTATTACACAGCTGTGTCTTCTGGATGC	800	QY	1745	GGCTGGGGCCCCCTGTCTGTGTTCTCAGCAGAAACCGGTGGCGGCTGGTGGGGCCCGTCC	1804
Db	471	GAATGAGTGGGCGCTGGAGCTGTCTCAGAGCTATTACACAGCTGTGTCTTCTGGATGC	530	Db	1551	GGCTGGGGCCCCCTGTCTGTGTTCTCAGCAGAAACCGGTGGCGGCTGGTGGGGCCCGTCC	1610
QY	801	GGGCTTCTCAGCCCTGCGAGAGCTCGGGCTCTTCTTCACTGGTACGACTCGCTTAC	860	QY	1805	ACCTGACCCGAGAGAGGGCGCTTTGGCTCAGCTTCGSGGAGACTGCGCTGTCTCTCA	1864
Db	531	GGGCTTCTCAGCCCTGCGAGAGCTCGGGCTCTTCTTCACTGGTACGACTCGCTTAC	590	Db	1611	ACCTGACCCGAGAGAGGGCGCTTTGGCTCAGCTTCGSGGAGACTGCGCTGTCTCTCA	1670
QY	861	TGGGGTCCCGGCCAGCAGCTGCGCTGGCTTTCGAGAGGCGAGCGTCTCTTCAACAT	920	QY	1865	TCGCTGCGCTTCATTCCAGGAGGACGAGCCCGCGGCTGGCTTCCAGAGAGGGGCGACTACA	1924
Db	591	TGGGGTCCCGGCCAGCAGCTGCGCTGGCTTTCGAGAGGCGAGCGTCTCTTCAACAT	650	Db	1671	TCGCTGCGCTTCATTCCAGGAGGACGAGCCCGCGGCTGGCTTCCAGAGAGGGGCGACTACA	1730
QY	921	CGGTGCCCTCCACACGAGATTGGGCGCGCCAGGACCGCTCTCTGCAACCGAGGAGTCCCG	980	QY	1925	TTGTGTCACTGAAATGGGCGAGCCATGCAAGTGTGTGAGACACGCGGAGGTGCTGACGAGC	1984
Db	651	CGGTGCCCTCCACACGAGATTGGGCGCGCCAGGACCGCTCTCTGCAACCGAGGAGTCCCG	710	Db	1731	TTGTGTCACTGAAATGGGCGAGCCATGCAAGTGTGTGAGACACGCGGAGGTGCTGACGAGC	1790
QY	981	CCGGCTATGAGGCTTTCAGAGGCGCTGGGGCTTTCAGCTCTCTGAGGAGAACTT	1040	QY	1985	TGAAGGCTCGGCGAGAGCGGCGCCAGCTGTCAGAGTGTGTGCTGCTGCTGCCAGCTCTA	2044
Db	711	CCGGCTATGAGGCTTTCAGAGGCGCTGGGGCTTTCAGCTCTCTGAGGAGAACTT	770	Db	1791	TGAAGGCTCGGCGAGAGCGGCGCCAGCTGTCAGAGTGTGTGCTGCTGCTGCCAGCTCTA	1850
QY	1041	CTCCATGCGCGAGCCAGACATGAGCGTGGCTGCTCTGCGCACTGGAGCAGCTCAT	1100	QY	2045	GACTGCCCAAGCTTTGGGGACCCGCGCGCTCTGCTGGGCCCCCAGGGGCTTCTAAGGA	2104
Db	771	CTCCATGCGCGAGCCAGACATGAGCGTGGCTGCTCTGCGCACTGGAGCAGCTCAT	830	Db	1851	GACTGCCCAAGCTTTGGGGACCCGCGCGCTCTGCTGGGCCCCCAGGGGCTTCTAAGGA	1910
QY	1101	GATGGCCAGGCCAGGAATGTGTGTTGAGGGCTCTCAACCTGCTCCATGGCCCC	1160	QY	2105	GCACAGGAGCAGTGTGTGCAAGACCCCGGCACTCAGTGGGCGCAGTCCCGGGCCCTCC	2164
Db	831	GATGGCCAGGCCAGGAATGTGTGTTGAGGGCTCTCAACCTGCTCCATGGCCCC	890	Db	1911	GCACAGGAGCAGTGTGTGCAAGACCCCGGCACTCAGTGGGCGCAGTCCCGGGCCCTCC	1970
QY	1161	CCAGACTGCTGCGCCAGCTGCGCTGGCGCAGGAGCGCGCAGGTGCGAGCCGAGTA	1220	QY	2165	TCAACTGGAGCCCAAAGGCGCAGCAGGCGCAAGACTGGAGGCTGCCCCAGCCTGTGCC	2224
Db	891	CCAGACTGCTGCGCCAGCTGCGCTGGCGCAGGAGCGCGCAGGTGCGAGCCGAGTA	950	Db	1971	TCAACTGGAGCCCAAAGGCGCAGCAGGCGCAAGACTGGAGGCTGCCCCAGCCTGTGCC	2030
QY	1221	CAGCTAGTGACCGGACCATGGCCAGCCAGCCCGTCCAGACTACGTGCTGTCTCTTG	1280	QY	2225	CAGTGAAGCAGCTTCGCGCTCATCTTTGAAGACCCAGGGTGGCGCTGAGGGCGAGAT	2284
Db	951	CAGCTAGTGACCGGACCATGGCCAGCCAGCCCGTCCAGACTACGTGCTGTCTCTTG	1010	Db	2031	CAGTGAAGCAGCTTCGCGCTCATCTTTGAAGACCCAGGGTGGCGCTGAGGGCGAGAT	2090
QY	1281	GACTGCCCTGGTGATGTAAGGCGGAGTACTTCCGCTCCCTGGCCCACTACAGTAGC	1340	QY	2285	CCCTGACGCGCTCAGCGCTGGCTCCAGCTGGCAGCAAGCACACGATGCCCTTCCCGAC	2344
Db	1011	GACTGCCCTGGTGATGTAAGGCGGAGTACTTCCGCTCCCTGGCCCACTACAGTAGC	1070	Db	2091	CCCTGACG--CTCAGCGCTTGGCTTCCAGCTTGGCAGCAGCAGCAGCAGTATGCCCTTCCCGAC	2149
QY	1341	CATGGCCCTCTGCCAGCGCTCCC-----	1363	QY	2345	CCAGAGAACTCCGGGCAATGCTGTCCGCTCATGCTGAGGCTGCTCTGGGCACTG	2404
Db	1071	CATGGCCCTCTGCCAGCGCTCCC-----	1330	Db	2150	CCAGAGAACTCCGGGCAATGCTGTCCGCTCATGCTGAGGCTGCTCTGGGCACTG	2209
				QY	2405	CCTGCCCAATTAAGACTGGTCAGACCTGTCTGA	2437


```

Db      2210 CCTGCCCTTAAGACTGCTGACAGCTGTCTGA 2242
|||||
RESULT 6
LOCUS   AY082588                2088 bp mRNA linear PRI 12-NOV-2002
DEFINITION Homo sapiens rhophilin-1 mRNA, complete cds.
ACCESSION AY082588
VERSION   AY082588.1 GI:19697912
KEYWORDS
SOURCE    Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2088)
AUTHORS   Peck, J.W., Oberst, M., Bouker, K.B., Bowden, E. and Burbelo, P.D.
TITLE     The RhoA-binding protein, Rhophilin-2, Regulates Actin Cytoskeleton
          Organization
JOURNAL   J. Biol. Chem. 277 (46), 43924-43932 (2002)
PUBMED   12221077
REFERENCE 2 (bases 1 to 2088)
AUTHORS   Burbelo, P.D.
TITLE     Structure and function of Rhophilin homologs
JOURNAL   Unpublished
REFERENCE 3 (bases 1 to 2088)
AUTHORS   Burbelo, P.D.
TITLE     Direct Submission
JOURNAL   Submitted (06-MAR-2002) Dept. of Oncology, Georgetown University
          Medical Center, 3970 Reservoir Rd., N.W., Washington, DC 20007, USA
FEATURES
         Location/Qualifiers
         1..2088
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:2155941"
         1..2088
            /note="non-kinase effector protein of RhoA; RhoA-binding
            protein"
         /codon_start=1
         /product="rhophilin-1"
         /protein_id="AA189809.1"
         /db_xref="GI:19697913"
         /translation="MILREEDGAGAEESPRLOQDSLTQIQQLQSRRAQIHQOI
         DKELQMTGAENLRATNNRVETVALELSYVNSNLQLEKELEELSGVDPRHGS
         EAVTPIPLGKETKEDWSTPLKELISVHFEDGAYEAEIRLEALRQMRTPSR
         NESGLLTAYNOLCFDARFLTPARSLGLFFHWDSLTVGPAQORALAFKGSVLFP
         NIGALHTQIGARQDRSCTEGARAMEAFORAAAGAFSLURENFSAFSDMSAASICAL
         BOLMAQAQECVFEGSLSPASAPQDCLAQLRLAQEAQVAAYELVHRTWAQPPVHD
         YVPVSWTALVHVKAERYRSLSAHYHVMALCDGSRCPHLPVLPFRPPRAGSPQLCPP
         AATSGELTHQVFLQPTTSKPRGVLPQLEERROIGKAHLKRALIQBEALRLHA
         LCRVLRVDLLRAVISQTLQSLAKYAELEDREDDFCEAAEAPDIQPKTHQPEARMPR
         LSKQGEDIFHRLGLPLSVFSANRWLVGPHVLTFRGGFGLTLRGSPVLIAAVPG
         SQAAAAGLKEGDYTVSVNGQPCRWWRHAEVVTTELKAAGEAGASLQVWSLLPSRLPSL
         GDPRPVLIGPRLRSQREHCKTPASTWASPRLLNWSRAQQGKTGGCQPCAPVK
         PAPSSLKHPWP"
         ORIGIN
         Query Match 75.7%; Score 1869; DB 9; Length 2088;
         Best Local Similarity 96.3%; Pred. No. 5,4e-298;
         Matches 1954; Conservative 0; Mismatches 0; Indels 75; Gaps 1;

QY      321 GGGCTGTGACTCCCTGAGCGAGATCCAGTGGCGGCAGCTGCAGAGCCGAGGCCAGAT 380
Db      |||||||
QY      60 GGGCTGTGACTCCCTGAGCGAGATCCAGTGGCGGCAGCTGCAGAGCCGAGGCCAGAT 119
Db      |||||||
QY      381 TCACCAGAGATTGACAGGAGCTGCAGATGGGAGCGGCGCTGAGAACCTTACAGAGC 440
Db      |||||||
QY      120 TCACCAGAGATTGACAGGAGCTGCAGATGGGAGCGGCGCTGAGAACCTTACAGAGC 179
Db      |||||||
QY      441 CACCAGCAACACCGGGTGGAGAGAGACGGTCCGCTCGAGCTGAGCTACGTCAACTCCAA 500
Db      |||||||
QY      180 CACCAGCAACACCGGGTGGAGAGAGACGGTCCGCTCGAGCTGAGCTACGTCAACTCCAA 239
Db      |||||||

501 CCTGCAGCTGCTGAAGAGGAGAGCTGAGAGAGCTCAGCGTGGCGTGGAGACCTTGGCCGGCA 560
Db      |||||||
240 CCTGCAGCTGCTGAAGAGGAGAGCTGAGAGAGCTCAGCGTGGCGTGGAGACCTTGGCCGGCA 299
Db      |||||||
561 TGGGAGCGAAGCTGTCACTGTCCCATGATCCCTTGGGCTTGAAGAGACCAAGAGCT 620
Db      |||||||
300 TGGGAGCGAAGCTGTCACTGTCCCATGATCCCTTGGGCTTGAAGAGACCAAGAGCT 359
Db      |||||||
621 GGACTGCTCTACACCGCTGAAGAGAGCTGATCTCAGTGCATTTGGAGAGAGCGGCCCTC 680
Db      |||||||
360 GGACTGCTCTACACCGCTGAAGAGAGCTGATCTCAGTGCATTTGGAGAGAGCGGCCCTC 419
Db      |||||||
681 CTACGAGGCGAGAAATCAGGAGCTGAGAGCCCTCGGCGAGGCGCATGCGGACCCCGAGCG 740
Db      |||||||
420 CTACGAGGCGAGAAATCAGGAGCTGAGAGCCCTCGGCGAGGCGCATGCGGACCCCGAGCG 479
Db      |||||||
741 GAATGAGTGGGCTTGGAGCTGCTCAGAGCTTATTAACACAGCTGTGCTTCTCGATGC 800
Db      |||||||
480 GAATGAGTGGGCTTGGAGCTGCTCAGAGCTTATTAACACAGCTGTGCTTCTCGATGC 539
Db      |||||||
801 GCGCTTCTCACCCTCGCAGGAGCTCGGGCTCTTTTCCACTGCTGAGCTCGCTTAC 860
Db      |||||||
540 GCGCTTCTCACCCTCGCAGGAGCTCGGGCTCTTTTCCACTGCTGAGCTCGCTTAC 599
Db      |||||||
861 TGGGCTCCCGGCCAGCAGCGTGGCTTCGAGAGAGGCGAGCGTTCTTCTCAACAT 920
Db      |||||||
600 TGGGCTCCCGGCCAGCAGCGTGGCTTCGAGAGAGGCGAGCGTTCTTCTCAACAT 659
Db      |||||||
921 CGGTGCTCTCAGAGAGAGTTGGGCGCGCAGACCGCTCTCGACCGAGGGTGCCTCG 980
Db      |||||||
660 CGGTGCTCTCAGAGAGAGTTGGGCGCGCAGACCGCTCTCGACCGAGGGTGCCTCG 719
Db      |||||||
981 CGGCTATATGAGGCGCTTCAGAGGCGCGTGGGCGCTTCACTGCTGAGGAGAACTT 1040
Db      |||||||
720 CGGCTATATGAGGCGCTTCAGAGGCGCGTGGGCGCTTCACTGCTGAGGAGAACTT 779
Db      |||||||
1041 CTCCATGCGCGAGCCAGACATGAGCGCTGCGTCTCTGCGCACTGGAGCAGCTCAT 1100
Db      |||||||
780 CTCCATGCGCGAGCCAGACATGAGCGCTGCGTCTCTGCGCACTGGAGCAGCTCAT 839
Db      |||||||
1101 GATGCGCCAGGCGCAGGAATGTGTTTGGGCGCTCTCAACACCTGCTCCATGGCCCC 1160
Db      |||||||
840 GATGCGCCAGGCGCAGGAATGTGTTTGGGCGCTCTCAACACCTGCTCCATGGCCCC 899
Db      |||||||
1161 CCAAGACTGCTGGCGCCAGCTGCGCTGCGCGAGAGCGCCCGCAGAGTGGCAGCGAGTA 1220
Db      |||||||
900 CCAAGACTGCTGGCGCCAGCTGCGCTGCGCGAGAGCGCCCGCAGAGTGGCAGCGAGTA 959
Db      |||||||
1221 CAGGCTAGTGACCGGACCATGGCCAGCCAGCCCGTCCACGACTACGTGCTCTCTCTG 1280
Db      |||||||
960 CAGGCTAGTGACCGGACCATGGCCAGCCAGCCCGTCCACGACTACGTGCTCTCTCTG 1019
Db      |||||||
1281 GACTGCTCTGTGATGTCAAGGCGGAGTACTTGGCTTCTGCGCCACTACGAGCTAGC 1340
Db      |||||||
1020 GACTGCTCTGTGATGTCAAGGCGGAGTACTTGGCTTCTGCGCCACTACGAGCTAGC 1079
Db      |||||||
1341 CATGCGCTCTGTGCGAGCGCTCCC----- 1363
Db      |||||||
1080 CATGCGCTCTGTGCGAGCGCTCCCAGTGGTGGCCCAACACTTGGCCATGTGTCGAAG 1139
Db      |||||||
1364 ----- 1385
Db      |||||||
1140 GCCCCCCCGCGCAGGCTCACAGCCTCTCTGTGTCCTCCAGCAGCAGCGAGAGAGTCCC 1199
Db      |||||||
1386 CACGCAAGAGAGAGTCTTCTGTGAGCCCCCACTCTCTTAAGCCCCAGGCGCTGTGCT 1445
Db      |||||||
1200 CACGCAAGAGAGTCTTCTGTGAGCCCCCACTCTCTTAAGCCCCAGGCGCTGTGCT 1259
Db      |||||||
1446 GCCCAGAGAGTGGAGAGAGCGCAGCTTGGCAAGGCAACCTGAGCGTGCATCTCT 1505
Db      |||||||
1260 GCCCAGAGAGTGGAGAGAGCGCAGCTTGGCAAGGCAACCTGAGCGTGCATCTCT 1319
Db      |||||||

```


through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 114 Row: k Column: 5
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 6680084.

FEATURES

source

1..2301
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="MGC:62337 IMAGE:5717786"
/tissue_type="Brain, mouse 12.5 dpc"
/clone_lib="NIH BMAP_FDO"
/lab_host="DH10B"
/note="Vector: pYX-ASC"
1..2301
/gene="Rhpn1"
/note="synonym: Rhophilin"
/db_xref="LocusID:14787"
/db_xref="MGI:1098783"
66..1421
/codon_start=1
/product="Rhpn1 protein"
/protein_id="AAH52010.1"
/db_xref="GI:30354357"
/db_xref="LocusID:14787"
/translation="MILEERPDDGGTGEBSRRPDDGSIKRGVGSFVNQPGQLQSHR
ARLHQISRELRTWRELYRATSNVTWRETVALELSYVNSLQLELAELSTSV
DVDPQEGEGITIMPLGKETKELDMATPKELISHFEGEDGSTFETFELELDLRQ
ATTPRDEAGLDMIIAOLFCFLOARFSPRSGLLFHWYDSLTGVPAQORALAF
EKSGLVFNIGALHTOIGARODCSCTECTNHAARAFORAAAGAFLLRNFTSHAPSPDMS
AASLMLEQLMIAQACRIFKGLLLPASATPDICPOLQLAQEAQVATYEYGLVHRAM
AQPPRYDYLPMASNTLVAHVAEHFCAALAHYAAMALCEHSPAKGELGARQEHVFPSTP
HEPLGPTLPQHPEDRRKLGGGAASHLVQSPAQGGPAAAGCGDSGTALPQGVILTARAR
G"

gene

CDS

misc_feature

189..404
/note="HRI; Region: Hrl repeat"
/db_xref="CDD:pfam02185"

misc_feature

408..966
/note="BR01; Region: BR01-like domain. This functionally
uncharacterised domain is found in a number of signal
transduction proteins, including Rhophilin and BR01"
/db_xref="CDD:pfam03097"

ORIGIN

Query Match 42.3%; Score 1043.8; DB 10; Length 2301;
Best Local Similarity 70.0%; Pred. No. 4.1e-162;
Matches 1550; Conservative 0; Mismatches 522; Indels 143; Gaps 5;

Qy	321	GGGCTGTGACTCCCTGACGAGATCCAGTGGCGCCAGTCGACAGCGCGCGCTGAGAACCTCTACAGAGC	380
Db	146	GGGCTATGGCTCTCTTTGTGAAACCAGCTGGCCAGTCGACAGTCAAGGGTCTGGCT	205
Qy	381	TCACAGCAGATTGACAAAGAGCTGCAGATGCGGACGGCGCTGAGAACCTCTACAGAGC	440
Db	206	CCACCAGCAGTACGACAGAGCTCCGAGTGGGACCGGTGCTGAAACCTCTACAGAGC	265
Qy	441	CACAGCAACACCGGTGAGAGAGCGTGGCCCTGGAGCTGAGCTACGTCACTCAA	500
Db	266	CACAGCAACACCTGGGTCCGAGAAACAGTGCACCTGGAGCTGAGCTACGTCACTCAA	325
Qy	501	CCTGACGTCTGAAGCAGAGCTGGAGGAGCTCAGCGGTGGCGTGGACCTGGCCGCA	560
Db	326	CCTGACGTCTGAAGAGAGCTGGCAGAGCTGAGCACCAAGCTAGTGTGGACAGCC	385
Qy	561	TGGAGCGAAGCTGTCTACTGTCCCATGATCCCTGGGCTGAAGAGACCAAGAGCT	620
Db	386	AGAGGCGAAGGATTACTATTCCATGATCCCTGGGCTGAAGAAACCAAGGAGTT	445
Qy	621	GGACTGGTCTACACCGCTGAAGAGCTGATCTCAGTGACCTTTGGAGGAGCGGCTC	680
Db	446	GGACTGGGCCACACCGCTGAAGAGCTGATCTCTGAGCACTTTGGAGAGGATGTACTCTC	505

Qy	681	CTACGAGGAGAGAAATCAGGGAGCTGGAGGCCCTGGCGAGGCCATGCGGACCCCGACCGC	740
Db	506	CTTTGAGACAGAGATCAGGAATAGAGACCTACGCGAGGCCACGAGGACTCCAGCAG	565
Qy	741	GAATGAGTGGGGCTGGAGCTGCTCAGCCTATTACAAACAGAGTGTGCTTCGTGATGC	800
Db	566	GGAGGAGGGGGCTGGATCTGCTCGCAGCCTACTATAGCAGCTCTGTTCTCTGGAGCG	625
Qy	801	GGCCTTCTCACCCCTGCCAGAGACCTCGGGCTCTTCTTCCATGGTACGACTCGCTTAC	860
Db	626	ACGTTTCTTCAGTCCCTCCAGGAGCCCGGGCTCTTCCATCTGGTACGACTCGCTTAC	685
Qy	861	TGGGGTCCGGGCCAGCAGCGTCCCTGGCTTCGAGAAAGGAGCAGCGCTTCTTTCACAT	920
Db	686	GGGGGTCCCGGCACAGCAGCGGGCTCTGGCTTCGAGAGGCGAGTACTCTGTTTACAT	745
Qy	921	CGGTGCTTCCACACGAGATTGGGGCGGCCAGGACCGCTCTCTGACCGAGGGTGCCTG	980
Db	746	TGGGGCACTCCACACGAGATTGGAGCACGCGAGGATTTGCTCTGCAACCAGGCGCAAA	805
Qy	981	CCGGCTATGGAGCCCTTCAGAGGGCGCTGGGGCTTCAGCCTCTCTGAGGAGAACTT	1040
Db	806	CCACGTGCTGAGGCTTTCCAGAGGGCTGCTGGGGCTTCAGACTCTCTGAGGAGAACTT	865
Qy	1041	CTCCATGCGCCAGACCATGAGCGCTGGCTCCCTCTGGGCACTGAGCAGACTCAT	1100
Db	866	CTCCACGCGCCAAAGCCAGACATGAGTGTGCTCTCTCTCTCCATGCTGAGCAGCTCAT	925
Qy	1101	GATGGCCAGGCCAGGAATGTGTTTGAAGGCTCTCAACCTGCTCCATGGGCCCC	1160
Db	926	GAITGCCAGGCAAGGAATGCATCTTCAAGGGCTTTTACTGCGACCTCTGCAACACC	985
Qy	1161	CCAGACTGCTGGCCAGCTGCGCTGGCGGAGGCGCCCGAGTGCAGCGCAGCGAGTA	1220
Db	986	TGACATCTTGTCCAGATCACTGACGTGGCCCGAGGAGCTGCCAGGTGCACTGAATA	1045
Qy	1221	CAGCTAGTGCACCGGACCATGGCCAGCCACCGCTCCAGACTACGTGCTCTCTCTG	1280
Db	1046	TGGGCTGTGCACCGGGCCATGGCCAGCCACCTGTCCGAGACTACCTGCTGCTCTCTG	1105
Qy	1281	GACTGCGCTGGTGTGATGTCGAAGCGGAGTACTTCCGCTCCCTGCGCCACTACCACTAGC	1340
Db	1106	GACTTAACCTGGCGCAGCTCAAGGCTGAGCCTTTGCGCCCTGGCCCACTATCATGAGC	1165
Qy	1341	CATGGCCCTCTGCGAGCGCTCCCGAGCGCCGAGGAGAGCTCCCGCAGCACGACGAGGT	1400
Db	1166	CATGGCCCTCTGCAAGAGCCACCCAG---CAAAGGTGAACTCGCAAGGAGGAGCAGT	1222
Qy	1401	CTTCTGACAGCCCCCACTCTCTTAAGCCCCGAGGCGCTGTGCTGCGCAGGAGCTGGA	1460
Db	1223	CTTC---CAGCCCTCCACTCCCATGAGCCCTGGGCCCCCAGCTGCGCAGCACCCAGA	1279
Qy	1461	GGAGCGCAGGAGCTTGGCAAGGCACACTGAAGCGTGCCATCTCTGGGGCAGGAGGAGC	1520
Db	1280	AGACCGCAGGAAGCTT-----GGAGGAGGC	1304
Qy	1521	GCTGCGCTGCAGCCCTGTGCGCCTCTGCGGAGGTGACCTGCTCGGGCTGTGAT	1580
Db	1305	GCTGCGCTTCAACACTTGTGAGTCTTGGCAGAGTCTGCGCAAGTGGACCTGCTGAGGTTG	1364
Qy	1581	CTCCAGACGCTGACGCTCCTCTGCGCAAGTATGCGAGCTGACCGTGAAGATGACTT	1640
Db	1365	GACTCAGGCACTGCGGCGCTCCCTGGCCAAAGTACTCAAGCTTCGAGCGCGAGGATGACTT	1424
Qy	1641	CTGTGAGGCTGCCGAGCGCCGAGCATCCAGCTTAAGACCCACAGAAAGCCAGAGGCCAG	1700
Db	1425	CTTCGAGGCTACGAGGCGCCGAGACATCCAAACCCAGACTCCACAGACCCAGAGGTGAG	1484
Qy	1701	GATGCCACGCTGTCCAGGGGAAGGGCTCTGACATCTTCCATCGGCTGGGGCCCTGTC	1760
Db	1485	GATGCCCTAGCTGTGCGAGGTGAAGTGAAGTACTTCTTCACTCGGCTGGGACCTTGTC	1544
Qy	1761	TGTGTTCTCAGCCAAAGAACCGGCTGGCGGCTGGTGGGGCCCGTCCACTGACCCGAGGAGA	1820

```
Db 1545 TGTGTTTTCACCAAGAACCGATGGCAGTGTGTAGGCGCGTGCAATGACTCGAGGAGA 1604
QY 1821 GGGCGGCTTTGGCTCAGCTTCGCGGAGACTCGCTGCTCCTCATGCTGCGCTCATTC 1880
Db 1605 GGGAAAGCTTCGGGTTTCAAGCTCGCGGAGACTCTCCAGTCTCTCATGCTGCTGCTGCC 1664
QY 1881 AGGGAGCCAGGCGCGG-----1897
Db 1665 TGGGGGTGAGGCTGAGGTAAGAGCCCTCTAGCCCGGGTGTGGAATCCTAGAGGACCGC 1724
QY 1898 -----CGGCTGGCTGAAGGAGG 1916
Db 1725 TAGATACTGCCACACCTCACCGATCCACATCTCTCGGAGTCGCTGGCTGAAGAGG 1784
QY 1917 CGACTACATTTGTTCAGTGAATGGGAGCAGCATGCGAGTGTGTGAGACACGCGAGGTGT 1976
Db 1785 CGACTACATCTGTCCTGTAACCGGAGCGGTCGCAAGTGTGTGAAGACCTTGAGGTGT 1844
QY 1977 GACGAGCTGAGGCTGCGGAGAGCGGGGCGCAGCTTGCAGTGTGTGCTGCTGCTGCC 2036
Db 1845 GACACAGCTGAGGAGCATGCGGCGAGGAAGCGTGTGAGCTGCAAGTGTGTCTATTGCTGCC 1904
QY 2037 CAGCTCTAGACTGCTCCAGCTTGGGGGACCGCGCGCGCTGCTGCTGGGCGCCAGGGGCT 2096
Db 1905 CAGCCTGAGCCACGTGGCAAGGACCCGAGCGGAGCAGCTGCTG-----1950
QY 2097 TCTAAGAGCCAGAGGAGCAATGTTGCAAGACCCCGCATCCAGCTGGGCGAGTCCCG 2156
Db 1951 -----TGGAAACAGAGAGTGTGTTTGTAGACACCGATGCCACACGTACCGCCCTG 2006
QY 2157 GCCCTCTCACTGGAGCCGAAAGCCAGGAGGCAAGACTGAGGCTGCGCCCGACCC 2216
Db 2007 GCCCATCTTGGCTGGAGCCGCAAGATAAACAAGGCAAGACTGGAAGCCACCTGACCC 2066
QY 2217 CTGTGCCCCAGTGAAGCAGCTCCCGCTCACTCTTGAAGACACCCAGGCTGCGCTGAGG 2276
Db 2067 CTGCACCAACAGGAATGTGTACCTGCCCCCTGAGCTGGAACACCCCGTATCCCTGTGTG 2126
QY 2277 GCGAGGATCCTGACGCGCTCAGCCCTGAGCTTCCAGCTGGCAGCAAGACCGAGCATGCC 2336
Db 2127 TCAGGATGGCTGTGGGGGCCATCTCTATGTATCATCACCTTGGAACTCAGCTCCAATAG 2186
QY 2337 CTCCCCACCCAGAGACCTCCGGGAATGCTGCTCCGCTCATGCTGAGGCTGCTGCTG 2396
Db 2187 CGAGATGCTCTAGCGGGTCTGTCTCTACTCTTGGGGACTATCAGCTGCGCCCACT 2246
QY 2397 GGCACCTGCTGCCCATTAAGACTGGTTCAGACCTGTCTGAAAAAAGGAGGAGGAGG 2451
Db 2247 TGCTTAGAGATGTCTATTAAAAATGCTGTGATGAAAAAAGGAGGAGGAGGAGGAGG 2301
```

RESULT 9

```
BD183400
LOCUS BD183400 4729 bp DNA linear PAT 17-JUN-2003
DEFINITION Novel genes and proteins encoded by the genes.
ACCESSION BD183400
VERSION BD183400.1 GI:31875600
KEYWORDS JP 2002345492-A/113.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1. (bases 1 to 4729)

Obara,O., Nagase,T. and Nakajima,D.

Novel genes and proteins encoded by the genes

Patent: JP 2002345492-A 113 03-DEC-2002;

KAZUSA DNA RESEARCH INSTITUTE

OS Homo sapiens (human)

PN JP 2002345492-A/113

PD 03-DEC-2002

PF 26-FEB-2002 JP 2002049009

PI OSAMU OHARA, TAKAHIRO NAGASE, DAISUKE NAKAJIMA

```
PC C12N15/09,C07K14/47//A61K31/711,A61K38/00,A61K48/00,A61P25/00,
PC A61P25/14,
PC A61P25/18,A61P35/00,C12N15/00,A61K37/02
CC Novel genes and proteins encoded by the genes FH Key
Location/Qualifiers
FT CDS (518)..(1540).
FEATURES
source Location/Qualifiers
1..4729
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
```

ORIGIN

```
Query Match 38.6%; Score 952.8; DB 6; Length 4729;
Best Local Similarity 93.2%; Pred. No. 3.5e-147;
Matches 1018; Conservative 0; Mismatches 72; Indels 2; Gaps 2;

QY 1346 CCCTCTGCGAGCGCTCCCTCAGGACCGGAGGAGAGCTCCCCACGACGAGCAGGTCTTCC 1405
Db 2235 CCCCGTGGCGGCACCCCCCAAGAAAGTGGCTGTGATGAGCCCCACAGCCCTGGCGTTGC 2294
QY 1406 TGCAGCCCCCACCCTCTTAAGCCCCGAGGCCCTGTGTCGCGCAGGAGCTGGAGGAGC 1465
Db 2295 CCATCTCTTCTGCCACGTTCCAGGGCCCAATGTTGTGACATCCAGTGC 2354
QY 1466 GAGCAGCTTTGGCAAGCACACCTGAAGCGTGCATCTCTGGGCGAGGAGGCGCTGC 1525
Db 2355 CCGCGGTGC-AGGCAAGGCACACCTGAAGCGTGCATCTCTGGGCGAGGAGGCGCTGC 2413
QY 1526 GCTGTGACGCGCTGTGCGCGCTCTGCGCGAGGTGGACCTGCTTCGGGCTGTGATCTCC 1585
Db 2414 GCTGTGACGCGCTGTGCGCGCTCTGCGCGAGGTGGACCTGCTTCGGGCTGTGATCTCC 2473
QY 1586 AGAGCTCTCAGCGCTCACTGCGCAAGTATGCGGAGCTCGACCGTGAGGATGACTTCTCTG 1645
Db 2474 AGAGCTCTCAGCGCTCACTGCGCAAGTATGCGGAGCTCGACCGTGAGGATGACTTCTCTG 2533
QY 1646 AGGTCTGCGAGCGCCCGGACATCCAGCTTAAGACCCACAGAACCCAGAGGCCAGGATGC 1705
Db 2534 AGGTCTGCGAGCGCCCGGACATCCAGCTTAAGACCCACAGAACCCAGAGGCCAGGATGC 2593
QY 1706 CAGGCTGTCTCCAGGAGGAGGCGCTGACATCTTCATTCGCTGGGCGCCCTGTCTGTGT 1765
Db 2594 CAGGCTGTCTCCAGGAGGAGGCGCTGACATCTTCATTCGCTGGGCGCCCTGTCTGTGT 2653
QY 1766 TCTCAGCAAGAACCGGTGGCGCTGGTGGGCGCTGACATCTTCATTCGCTGGGCGCCCTGTCTGTGT 1825
Db 2654 TCTCAGCAAGAACCGGTGGCGCTGGTGGGCGCTGACATCTTCATTCGCTGGGCGCGG 2713
QY 1826 GCTTTGGCTCACGCTTTCGGGAGAGCTCGCTGTCTCATCTGCTGCGCTCATTCAGGGA 1885
Db 2714 GCTTTGGCTCACGCTTTCGGGAGAGCTCGCTGTCTCATCTGCTGCGCTCATTCAGGGA 2773
QY 1886 GCGAGGCGCGCGCTGGCTTGAAGAGGCGGAGTACATTTGTCTAGTGAATGGGAGC 1945
Db 2774 GCGAGGCGCGCGCTGGCTTGAAGAGGCGGAGTACATTTGTCTAGTGAATGGGAGC 2833
QY 1946 CATGCAAGTGTGAGACACCGGAGGTGTGACGAGCTCAAGGCTCGCGGAGAGGCGG 2005
Db 2834 CATGCAAGTGTGAGACACCGGAGGTGTGACGAGCTCAAGGCTCGCGGAGAGGCGG 2893
QY 2006 GCGCCAGCTGCAAGTGTGTGCTGTGCTGCCAGCTCTAGACTGCCAGCTTGGGGGACC 2065
Db 2894 GCGCCAGCTGCAAGTGTGTGCTGTGCTGCCAGCTCTAGACTGCCAGCTTGGGGGACC 2953
QY 2066 GCGCCAGCTGCTGCTGCTGGGCGGCTTCTAAGAGGCGGAGGAGGAGTGGTGA 2125
Db 2954 GCGCCAGCTGCTGCTGCTGGGCGGCTTCTAAGAGGCGGAGGAGGAGTGGTGA 3013
QY 2126 AGACCCCGGATCCACGCTGGGCGAGTCCCGGCGCCCTCTCACTGAGGCGGAGGAGGCG 2185
Db 3014 AGACCCCGGATCCACGCTGGGCGAGTCCCGGCGCCCTCTCACTGAGGCGGAGGAGGCG 3073
```

Qy	2186	AGCAGGCACAGACTGGAGGCTGCCCCAGGCCCTGTGCCCCAGTGAAGCCAGCTCCGCCCT	2245
Db	3074	AGCAGGCACAGACTGGAGGCTGCCCCAGGCCCTGTGCCCCAGTGAAGCCAGCTCCGCCCT	3133
Qy	2246	CATCCTTTGAAGCACCCAGGCTGGCCGTGAGGGCCAGGATCCCTGCACGCCCTCAGCCCTG	2305
Db	3134	CATCCTTTGAAGCACCCAGGCTGGCCGTGAGGGCCAGGATCCCTGCACGCCCTCAGCCCTG	3192
Qy	2306	GCTCCAGCTGGCAGCAAGCACCCAGCATGCCCTCCCAACCAGAGGACCTCCGGGCAATG	2365
Db	3193	GCTCCAGCTGGCAGCAAGCACCCAGCATGCCCTCCCAACCAGAGGACCTCCGGGCAATG	3252
Qy	2366	CCTGTCCCGCTCATGTGGAGGCTGCCTCGGGCACCTGCTGCCATTAAAGACTGGTC	2425
Db	3253	CCTGTCCCGCTCATGTGGAGGCTGCCTCGGGCACCTGCTGCCATTAAAGACTGGTC	3312
Qy	2426	AGACCTGTCTGA	2437
Db	3313	AGACCTGTCTGA	3324

[illegible]

13333333

REFERENCES

AUTHORS:

FREE

ETIT

JOURNAL

MEDT.TN

INTRODUCTION

PUBMED

OPTION


```

RESULT 12
BC036447
LOCUS
DEFINITION
Homo sapiens rhophilin, rho GTPase binding protein 2, mRNA (cDNA
clone MGC:26769 IMAGE:4830913), complete cds.
ACCESSION
BC036447
VERSION
BC036447.2
KEYWORDS
GI:34784999
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
REFERENCE
1. .3516
/notes="Vector: pBluescript"
/genes="RHPN2"
/notes="synonym: p76RBE"
/db_xref="LocusID:85415"
37. .2097
/codon_start=1
/product="rhophilin-like protein"
/protein_id="AAH36447.1"
/db_xref="GI:22209019"
/db_xref="LocusID:85415"
/translation="MTDALPAPQPLEKNDGYFRKGNPLAQTRSKLQNRQAALN
QQILKAVRMTGAENLKVATNSKVRQVLELSFVNSDIQMLEEGLNLSVGVYQ
NTEAETPLIPLGLKXETDQVFAVVKDFILEHSGEDGYLYVEIDADMLRQACT
PSDEAGVELLMTYFQLGQVSRFFPTOMGLLFTWYDSLTVGPVSOQLLEKAS
VLNTGALYQIGTRCDRQTOAGLESADAFQAAVNLKDTFTHTPSYDMSAPML
SVLVKMLAQAESVFEKISLPGRNEFPMVKVQAEAAKGVYQQLHAAMSQAPVK
ENIPYSIACVKAHYAALAHYFTAILLDHVKPQPTDLDHOEKLQCLSDHPMG
LTLATLKNQDQRRQLKSHLRMAHHEESVREASLKKLSIEVLQKLVCAQERS
RLTYAQHQBEDDLNLDAPSVAKTEQEVDIILPQFSKLTVDTPFKLGLPVSFAS
KWTTPRSIRFTAEGDGLGFLRGNAPVQVHFLDPYCSASVAGAREGYIVSLVDC
KWLTLSEVMKLLKSFGEDETEMKVSLDSTSMHNKSATYSVMQKTYGMICLAIDD
DKDKTKIKSLKSLFSLSGTNRKQKASTLCLPSVGAARPQVKKLPSFSLNSD
SSWY"
misc_feature
163. .354
/notes="Hr1; Region: Hr1 repeat"
/db_xref="CDD:pfam02185"
367. 825
/notes="BR01; Region: BR01-like domain. This functionally
uncharacterised domain is found in a number of signal
transduction proteins, including Rhophilin and BR01"
/db_xref="CDD:pfam03097"
misc_feature
1576. .1809
/notes="PD2; Region: Domain present in PSD-95, Dlg, and
ZO-1/2. Also called DHR (Dlg homologous region) or GLGF
(relatively well conserved tetrapeptide in these domains).
Some PDZs have been shown to bind C-terminal polypeptides"
/db_xref="CDD:smart00228"
ORIGIN
Query Match 20.1%; Score 496.4; DB 9; Length 3516;
Best Local Similarity 56.8%; Pred. No. 5.2e-72;
Matches 977; Conservative 0; Mismatches 731; Indels 12; Gaps 3;
Qy 321 GGGCTGTGACTCCCTGACCGAGATCCAGTGGCGGCGAGCGAGCGGCCAGAT 380
Db 105 GGGCTGTGACTCCCTGACCGAGATCCAGTGGCGGCGAGCGGCCAGAT 164
Qy 381 TCACCGACGAGATTGACAAAGGAGTGCAGATCGGACGGGGCGCTGAGAACCTCTACAGAGC 440
Db 165 GAATCAGCAGATCTCTGAAAGCGCGTGGCGGATGAGGACCGGAGCGGAAACCTCTGAAAGT 224
Qy 441 CACGAGCAACACCGGCTGAGAGAGCGTGGCGGCTGAGCTGAGCTACGCTCAACTCAA 500
Db 225 GGCCACAAACTCAAAGGTGGCGGAGCAAGTGGCGGCTGAGCTGAGCTTCTGCAACTCAGA 284
Qy 501 CCTGAGCTGTCTGAAAGGAGGAGTGGAGGAGCTCAGCGGCTGGCGGTGGACCCCTGGCCGCA 560
Db 285 CCTGAGCTGTCTGAAAGGAGGAGTGGAGGAGCTCAGCGGCTGGCGGTGGACCCCTGGCGCA 344
Qy 561 TGGGAGGAGAGCTGTCTGCTGCCCATGATCCCTGGGCTTGAAGAGACCAAGGAGCT 620
Db 345 CACAGAGGAGGAGCAATTTACGATTCCTTCCTTCCTGGCTGAGGAAAGCAAGCAAGT 404
Qy 621 GGACTGGTCTACACCGCTGAAGGAGCTGATCTCAGTGCACCTTTTGGAGGACCGCGCTC 680
Db 405 CGACTTTGAGTCTCTCTCAAGGATTTTATCTGAGCAATTTACAGTGAAGTGGCTATTT 464
Qy 681 CTACGAGGAGCAAAATCAGGAGGAGTGGAGGCGCTCGGCGAGGCCATCGGACCCGAGCG 740
Db 465 ATATGAAGATGAAATTCAGATCTTATGATCTGAGACCAAGTCTGTCGAGCGCTAGCG 524
Qy 741 GAATGAGTGGGCGCTGGAGCTGTCTACAGCCCTATTACACAGCTGTCTTCTTGATGC 800

```


Db 525 GATGAGGCGGGTGGAACTGCTGATGACATACTTCATCCAGCTGGGCTTTGTGAGAG 584
QY 801 GCGCTTCTCCACCCCTGCGAGGAGCCTCGGGCTCTTTCTTCCACTGTGACGACTCGCTTAC 860
Db 585 TCGATTCTTCCCGCCACACGCGCAGATGGACTCTCTGTTTCACTGGTATGACTCTCTCAC 644
QY 861 TGGGGTCCCGCCAGCAGCGTGCCTGCGCTTTCAGAAAGGGAGCGTTCTCTTCAACAT 920
Db 645 CGGGTTCGGTFCAGCAGCAGAACCTGCTGTGGAGAGGCGAGTGTCTCTGTTCAACAC 704
QY 921 CGGTGCGCTTCCACACGACGATTTGGGCGCGCGCAGAGACCGCTCTCCACGAGGCTGCCCG 980
Db 705 TGGGGCGCTTACACCCAGATTGGGACCCGGGTGTGATCGGAGAGCGAGCTGGGCTGA 764
QY 981 CGCGCTATGGAGGCTTCCAGAGGCGCGCTGGGGCTTTCAGCGCTTCCAGGAGGAACTT 1040
Db 765 GAGTCCCATAGATGCTTTCAGAGAGCGCGAGGGTTTTAAATTACCTGAAGACACATT 824
QY 1041 CTCCTATGCGCGAGCCAGACATAGCGCTGCGTCCCTCTGCGCACTGGAGCGCTCAT 1100
Db 825 TACCCATACTCCAAGTTACGACATGAGCCCTGCCATGCTCAGCGTCTCGTCAAAATGAT 884
QY 1101 GATGCGCCAGGCGCAGGAATGTGTTGAGGCGCTCTCACCACCTGCTCCATGCGCCC 1160
Db 885 GCTTGCAACAGCCCAAGAAAGCGTGTGAGAAATCAGCCTTCTCTG-----GATCG 938
QY 1161 CAAAGACTGCTGCGCCAGCTGCGCTGCGCGAGGAGCGCGCCAGGTGGCAGCGAGTA 1220
Db 939 GAATGAATTCCTATGCTGTGTAAGGTGCTCAGAGAGCTGTAAAGTGGGAGAGTCTA 998
QY 1221 CAGGTAGTGACCGGACCATGCGCCAGCACCCTGCAGACTACGTGCGCTCTCTG 1280
Db 999 CAAACAGTACACGACGACCATGAGCAGCGCGCGGTGAAGAGAAATCCCTTACTCTCTG 1058
QY 1281 GACTGCCCTGCTGATGTCAAGCGCGAGTACTTCCGCTCCCTGCGCCCACTACCGATGAC 1340
Db 1059 GCGCAGCTTAGCTCGTGAAGGCCCACTACGCGGCGCTTGGCCCACTTCTCCTG 1118
QY 1341 CATGCGCTCTGCGAGCTCCCGCAGCAGCAGGAGAGCTCCCGCAGCAGCAGCAGT 1400
Db 1119 CATCTCTCATCGACCAACAGGTGAAGCCAGCGACGGATCTGGACCAACAGGAGAGTG 1178
QY 1401 CTCTCTGAGC-----CCCCAGCTCTTAAGCCCGGAGGCGCTGTGC--TGCCGAGGA 1454
Db 1179 CCTGTCCAGCTCTAGCACCACATCCAGAGGGGCTGACACCCCTTGGCCACACTGAAGA 1238
QY 1455 GCTGAGGAGCGCAGCGAGCTTGGCAAGCGCACCTGAAGCGTGCATCTCGGGCAGGA 1514
Db 1239 TGATCAGCAGCGCCGACAGCTGGGAAGTCCCACTTGGCAGAGCCATGGCTCATCAGA 1298
QY 1515 GAGGCGCTGCGGTGACGCGCTGTGCGCGCTCTGCGCGAGGTGGACCTGCTTCGGGC 1574
Db 1299 GGAGTCGTGCGGAGGCGCAGCTCTGCAAGAGCTGCGGAGCATTTAGGTGCTACAGA 1358
QY 1575 TGTGATCTCCAGAGCTGCGAGCGTCTACTGCGCAAGTATGCGGAGCTGACCGGTGAGA 1634
Db 1359 GGTGCTGTGTCCGACAGGAACCGTCCCGGTCACTGCGCGCTACGTCACGCCACAGGAGGA 1418
QY 1635 TGACTTCTGTGAGGTGCGGAGGCGCCGACATCCAGCTTAAGACCCACAGAGCCAGA 1694
Db 1419 TGACTGCTGAACCTGATCGAGCGCCCGCCAGTGTGTTGCTTAACTGAGCAAGAGTTGA 1478
QY 1695 GGCCAGATGACACGCTGTCCAGGGGAAGGGGCTTGACATCTTCCATCGGCTGGGGCC 1754
Db 1479 CATTATATTGCCCGAGTTCTCAAGCTGACAGTCAGGACTTCTTCCAGAGAGCTGGGCC 1538
QY 1755 CTGTCTGTGTTCTCAGCCAGAACCCGCTGGCGGTGTTGGGCGCCCGTCCACCTGACCCG 1814
Db 1539 CTTATCTGTGTTTTCGGGTAAACAGCGGTGAGCGCTCTCTCGAAGCATCCGCTTCACTGC 1598
QY 1815 AGGAGAGCGCGCTTTGGCTTCAAGCTTCGGGAGACTCGCTGTCTCTCATCGCTGCCGT 1874

Db 1599 AGAAGAAGGGGACTTGGGGTTCACTTGAGAGGGAACGCCCGCTTCAAGTTCACTTCT 1658
QY 1875 CATTCAGGAGGCCAGCGCGCTGGCTGCTGAAGGAGGCGACTCATATTGTGTCACT 1934
Db 1659 GGATCTTACTCTGCTGCTGCTGCGTGGCAGGAGCCCGGAGAGAGATTATTTGTCTCAT 1718
QY 1935 GAATGGGAGCCATGCAAGTGGTGGAGACACCGGAGGTGGTACGAGCTGAAGGCTGC 1994
Db 1719 TCAGCTTGTGATTGTAAGTGGCTGACGCTGAGTGAGGTATGAAGCTCTGAAGAGCTT 1778
QY 1995 GGGAGAGGCGCGCGCCAGCCTGACGCTGAGTGGTGTGCTGCTG 2034
Db 1779 TGGCAGGAGCAGATCGAGATGAAGTCTGTAGCCTCTG 1818

RESULT 13

HSA347750

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

HSA347750 2061 bp mRNA linear PRI 06-JAN-2003
Homo sapiens mRNA for p76RBE protein.
AJ347750
AJ347750.1 GI:23304861
p76RBE protein.
Homo sapiens (human)

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 Mitescu, H., Steuve, S., Savonet, V., Degraef, C., Mellor, H.,
Dumont, J.E., Maenhaut, C. and Pirson, I.
Identification and characterization of a novel activated RhoB
binding protein containing a PDZ domain whose expression is
specifically modulated in thyroid cells by cAMP
Eur. J. Biochem. 269 (24), 6241-6249 (2002)

2 (bases 1 to 2061)
12473120
Pirson, I.
Direct Submission
Submitted (20-SEP-2001) Pirson I., Interdisciplinary Institute
IRIBHN, Université Libre de Bruxelles, 1070, Brussels, BELGIUM
Location/Qualifiers
1..2061
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..2061
/function="putative RhoB effector"
/codon_start=1
/product="p76RBE protein"
/protein_id="CAC87939.1"
/db_xref="GI:23304862"
/db_xref="GOA:Q8IUC4"
/db_xref="SPTREMBL:Q8IUC4"
/translation="MTDALLPAAPQPLEKENDGYFRKGNPLAQTGRSKLQNRALN
QQILKAVRMRTAGNLLKVNATNSQVRLSELSFVNSDLQMLKEELNLSVGVYQ
NTEAFATPLTGLKETDQVDFVNLKDFILEHSEYDGLYEIDLEADMLRACRT
PSRDEAGVELIMTYFIOLGFVRSFPFPTOMGLFTWYDSLTVGPVSQOQLLEKAS
VLFTNGALYTGITGRCDTQAGLESAIDAFORAAVNLKDTFTHTPSVDMSPAML
SVLVKMLAQAESVFKEISLPGIRNFFMLVKVAQEAARKVEYVQQLHAAMSQAPVK
ENIPYSWASLACVKAHYAALHYFTAILLDHQVKFGTDLDHOKELCSQLYDHMEG
LTPATLNQDQORQLKSHLRAMAHHEESVREASLCKLRSLEVLQKVLCAOERS
RLTYAQEEDDLNLIDAFSVVAKTEQVDIILPQSKLTVTDFQKGLPSVFSAN
KRWTPRSIRFTABEGDGLFTRGNAPOVQVHFLDPYCSASVAGAREGDIYVSIQLVDC
KWLFTLSVMKLLKSGFDELEMKVVSILDTSSMHNKSATYVSGMQKTYSMICLAIDD
DDKTDTKXISKSLSLFSLSWGNTKQRQKASATLCLPSVGAARPVKKLLPSPFLNSD
SSWT"

ORIGIN

Query Match 20.0%; Score 494.8; DB 9; Length 2061;
Best Local Similarity 56.7%; Pred. No. 1e-71;
Matches 976; Conservative 0; Mismatches 732; Indels 12; Gaps 3;
QY 321 GGGCTGTGACTTCCCTGACGCGAGATCCAGTGGCGGCAGCTGCAGAGCGCGAGGCCAGAT 380

```
Db      69  GGGCTGTAATCCCTTGCACAAACCGCGCGAGTAATTCAGAAATCAAGAGCTGCTTT 128
Qy      381 TCACGACGAGATGACAAAGGAGTGTGAGATGCGGACGGGCGCTGAGAACTCTACAGAGC 440
Db      129 GAATCAGCAGATCCTGAAAGCCGTGGGATGAGACCGGAGCGGAAACCTTCTGAAAGT 188
Qy      441 CACAGCAACAAACGGGTGAGAGACGCTCGCCCTGGAGCTGAGCTACGCTCAACTCCAA 500
Db      189 GGCACAAACTCAAGGTGCGGAGCAAGTGCAGCTGAGCTGAGCTTCTCAACTCAGA 248
Qy      501 CTTGAGCTCTCAAGAGGAGCTGAGGAGCTCAGCGGTGGGCTGAACTCTCGTGGGCTCTATCAGAA 308
Db      249 CTTGAGATGCTCAAGGAGAGCTGAGGAGCTGAGGAGCTGAACTCTCGTGGGCTCTATCAGAA 308
Qy      561 TGGAGCGAAGCTGTCACTGTCCCATGATCCCCCTGGGCTGAGAGAGACCAAGAGAGCT 620
Db      309 CACAGAGGAGGCAATACGATTCCTCTGATTCCTCTTGGGCTGAAAGGAAACGAAAGAGCT 368
Qy      621 GGAAGTGTACACCGCTGAAGAGCTGATCTCAGTGCATCTTGGAGAGACGCGGCTC 680
Db      369 CGACTTTGAGTGTCTCAAGGATTTATCTTGGAAATATACGTGAAGATGGCTATTT 428
Qy      681 CTAAGAGGCAAGAAATCAGGAGCTGAGGAGCTGCGGAGGCTGCGGAGGCTGCGGAGGCTGCGG 740
Db      429 ATATGAAGATGAATTCAGATCTTATGATCTGAGACAGCTTGTGCGAGCTAGCGG 488
Qy      741 GAATGAGTGGGCTGTGAGCTGTCTACAGGCTATTAACAACAGCTGTGCTTCTGAGTGC 800
Db      489 GGATGAGGCGGGGTGAACTGTGTATGATGATGATGATGATGATGATGATGATGATGATGAT 548
Qy      801 GCGCTTCTCAGCCCTGCGAGGAGCTGCGGCTCTTCTTCCACTGTGTACACTGCTTAC 860
Db      549 TCGAATTTCCCGGCGCACGCGAGATGGGACTCTCTGTTCACCTGTGTATGACTCTCTAC 608
Qy      861 TGGGGTCCCGGCGCAGCAGGCTGCGCTTGGGCTTCGAGAGGCGAGCGCTTCTCTCAACAT 920
Db      609 CGGGTTCCGCTCAGCCAGCAGAACCTGCTGTGAGAGAGCGGAGCTGCTGTCTCAACAC 668
Qy      921 CGGTGCGCTCAGCAGGATTTGGGCGCGCAGGACCGTCTCTGACCGAGGCTGCGCG 980
Db      669 TGGGGCGCTCTACACCCAGATTTGGGACCGGCTGTGATCGGACAGCGAGCTGGGCTGGA 728
Qy      981 CCGGCTATGAGGCTTTCAGAGGCGCTGCGGCTTTCAGGCTCTGAGGAGAGAACTT 1040
Db      729 GAGTGCCATGATGCTTTTCAGAGCGCGAGGAGGTTTAAATACCTTGAAGACACATT 788
Qy      1041 CTCCTATGCGCGGAGCCGAGACATGAGCGCTGCGTCCCTTGGGCACTGGAGAGCTCAT 1100
Db      789 TACCATACTCCAAATTAACACATGAGCCCTGCGCATGCTCAGCGTGTCTGCTCAAAATGAT 848
Qy      1101 GATGGCCGAGCCGAGGAAATGTGTGTTGAGGCTCTCACACCTGCTCCATGCGGCC 1160
Db      849 GCTGTGCAAGCCCAAGAAACGCTTTGAGAAATCAGGCTTCTCTG-----GGATCCG 902
Qy      1161 CCAAGACTGTGTCGCCAGCTGCGCTGCGGAGGAGGCGCGGAGGCTGAGGAGGAGTCA 1220
Db      903 GAATGAATTTCTATGCTGTGTGAGGTGGCTCAGAGGCTGTAAAGTGGAGAGGCTCA 962
Qy      1221 CAGGCTAGTGCACGGGCAATGCGGCCAGCCACCGCTCCAGACTACGTGCTGTCTCTG 1280
Db      963 CCAACAGCTACACGAGCCATGAGCCAGGCGCGGTGAAAGAGAAATCCCTTACTCTG 1022
Qy      1281 GACTGCCCTGCTGATGCTCAAGCGGAGTACTTCCGCTCCGTCGCGCCACTACCACTGAGC 1340
Db      1023 GGCAGCTTAGCTGTGAGGCGCCACCACTACGCGGCTTGGCCCACTTCTACTGTC 1082
Qy      1341 CATGGCCCTCTGCAACGGCTCCCGAGCGAGGAGAGCTCCCGCAGCAGCAGCAGGAGT 1400
Db      1083 CATCTCTCTCATCGACCAACAGGTGAGCCAGGACCGGATCTGGACCAACAGGAGAGTG 1142
Qy      1401 CTTCCCTGAGC-----CCCCACCTCTCTTAAGCCCGAGGCGCTGTGC--TGCCGAGGA 1454
```

```
Db      1143 CTTGTCCAGCTCTACGACCAACATGCGAGAGGCGCTGACACCCCTTGCCACACTGAGAA 1202
Qy      1455 GCTGGAGGAGCGCAGGAGCTTGGCAAGGCACACCTGAGAGCTGACCTCTGGGGCAGGA 1514
Db      1203 TGATCAGCAGCGCGCAGACAGCTGGGAGAGTCCCACCTTGGCAGAGCCATGCTCATCAGA 1262
Qy      1515 GGAGGCGCTGCGGCTGCAAGCGCTGTGCGCGCTCTCTGCGAGGTGAGACCTGCTTGGGGC 1574
Db      1263 GGAATCGGTGGGAGCGGAGCTCTGCAAGAGCTGCGGAGCATTTAGGTGCTTACAGAA 1322
Qy      1575 TGTGATCTCCAGAGCGTGCAGCGCTCACTGGCCAAAGTATGCGGAGCTTCACCGTGAAGA 1634
Db      1323 GGTGCTGTGTGCGCGCACAGGAACTCTCCCGCTCAGCTACGCTACGCGCAGCAGGAGGGA 1382
Qy      1635 TGAATTTCTGTGAGGCTGCGGAGCGCGGACATCCAGCTTAAGACCCACCAGAGCCAGA 1694
Db      1383 TGACCTGCTGAACCTGATCGACCGCCCCAGTGTGTTGCTTAAACTTGAGCAAGGTTGA 1442
Qy      1695 GGCCAGGATGCCACGCTGTCCAGGGAAGGGGCTGACATCTTCCATCGGCTGGGGCC 1754
Db      1443 CATTATATTGCCCCAGTCTTCCAAGCTGACAGTCAAGGCTTCTTCCAGAGCTGGGCC 1502
Qy      1755 CTTGTCTGTGTTCTAGCCAAAGAACCGGTGGCGCTGTGGGCGCGCTCCACCTGACCCG 1814
Db      1503 CTTATCTGTGTTTTCGCTAAACAAGCGGTGAGCGCTCTCTCGAAGCATCCGCTACTGC 1562
Qy      1815 AGGAGGCGCGCTTGGGCTCAGCTTCCGGGAGACTGCGCTGTCTCATCGCTGCGCT 1874
Db      1563 AGAAGAGGGGACTTGGGTTTCACTTGAGAGGAAACGCCCCGTTTTCAGGTTTCACTTCT 1622
Qy      1875 CATTCCAGGAGGAGCGCGCGCTGCGCTGAAAGGAGGCGGCTGAAAGGAGGCTTGTGTCAGT 1934
Db      1623 GGATCTCTACTGCTGCTGCTGCGTGGCAGGAGCGCGGAGGAGATTTATTTGCTCCAT 1682
Qy      1935 GAATGGCAGCCATCAGGTGGTGGAGACAGCGGAGGTGTCACGAGCTGAGGCTGAGGCTGC 1994
Db      1683 TCAGCTTGTGATTTAAAGTGGCTGACGCTGAGTGGGTTTATGAGCTGCTGAGAGCTT 1742
Qy      1995 GGAGAGGCGCGCGCGCAGCTGCGAGTGTGCTGCTGCTG 2034
Db      1743 TGGCGAGGAGCAGATCAGATGAAGTCTGTGAGGCTTCTG 1782
```

RESULT 14

AX835273 LOCUS AX835273 3500 bp DNA linear PAT 15-DEC-2003
DEFINITION Sequence 2397 from Patent EPI347046.

ACCESSION AX835273
VERSION AX835273.1 GI:39921408
KEYWORDS
SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1
AUTHORS Isogai, T., Sugiyama, T., Otsuki, T., Wakamatsu, A., Sato, H., Ishii, S., Yamamoto, J. I., Isono, Y., Hio, Y., Otsuka, K., Nagai, K., Irie, R., Tamechika, I., Seki, N., Yoshikawa, T., Otsuka, M., Nagahari, K. and Masuko, Y.

TITLE

Full-length cDNA sequences

JOURNAL

Patent: EP 1347046-A 2397 24-SEP-2003;

FEATURES

Location/Qualifiers (JP)
source
1..3500
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 20.0%; Score 494.8; DB 6; Length 3500;
Best Local Similarity 56.7%; Pred. No. 9.5e-72;
Matches 976; Conservative 0; Mismatches 732; Indels 12; Gaps 3;

COMMENT

NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.

FEATURES

source

1.3500
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UTER2006593"
/tissue_type="uterus"
/clone_lib="UTERU2"
/note="cloning vector: pME189FL3"

ORIGIN

Query Match 20.0%; Score 494.8; DB 9; Length 3500;

Best Local Similarity 56.7%; Pred. No. 9.5e-72;
Matches 976; Conservative 0; Mismatches 732; Indels 12; Gaps 3;

```
QY 321 GGGCTGTGACTCCCTGACGAGATCCAGTGGCGCCAGCTGCAGAGCGCGAGGGCCAGAT 380
DB 105 GGGCTGTATCCCTTGCACAAACCGCCGGAGTAATTCAGAAATCAAGAGCTGCTT 164
QY 381 TCACAGCAGATTGACAAAGAGCTGCAGATGCGGAGCGGGCGCTGAGAACCTTACAGAGC 440
DB 165 GAATCAGCAGATCTGAAAGCCATGCGGATGAGGACCGGAGCGGAAACCTTCTGAAAGT 224
QY 441 CACAGCAACACCGGCTGAGAGAGCGTCCGCTGGAGTACGCTACCTCAACTCAA 500
DB 225 GGCACAACTCAAGGTGCGGAGCAAGTGGCGCTGGAGTGGCTTCTGTAAGTCTGAA 284
QY 501 CTTGCACTGCTGAAGGAGAGCTGGAGAGCTCAGCGGTGGCGTGGACCCCTGGCGGCA 560
DB 285 CTTGCACTGCTCAAGGAGAGCTGGAGGAGCTGAACATCTCGGTGGCGCTTATCAGAA 344
QY 561 TGGAGCGAGCTGCTACTGTCCCTCATGATCCCTGGCGCTGAGAGAGCAGAGAGCT 620
DB 345 CACAGAGGAGGCAATTTACGATTCCTGATTCCTCTTTGGCGTGAAGGAAACGAAAGC 404
QY 621 GGAATGCTTACACCGCTGAGGAGCTGATCTCAGTGCATCTTGGAGAGAGCGCGCTC 680
DB 405 CGACTTTGAGTCTGCTCAGGAGTTTATCTGGAACATTAAGTGAAGATGCTATTT 464
QY 681 CTACAGGCGAAGATCAGGAGCTGGAGGCTCGGAGGCCATCGGAGCCATCGGACCCCGCG 740
DB 465 ATATGAAGATGAATTCAGATCTTATGATCTGAGACAAAGCTTGTGGAGCGCTAGCG 524
QY 741 GAATGAGTGGGCTCGAGCTGCTCAGAGCTTATACACAGAGTGTGCTTCTGGATGC 800
DB 525 GGATGAGGCGGGGTGAACTGTGATGATGATCTTATCAGCTGGGCTTTGTCGAGAG 584
QY 801 GCGCTTCTCACCCCTGAGAGGCTCGGCTCTTCTTCCACTGATGAGTACGCTGCTTAC 860
DB 585 TCGATTTTCCGCGCCACACGAGATGGGACTCTGTTTACCTGGTATGATCTCTCAC 644
QY 861 TGGGCTTCCGCGCCAGAGCTGCTGCGCTTTCGAGAGGAGCGCTTCTTCAACAT 920
DB 645 CGGGTTCCTGATCGAGCAGAGACCTTGTCTGGAGAGGCGGAGTGTCTTGTTCACAC 704
QY 921 CGGTGCTTCCAGAGAGTGGGCGCGGAGAGCGCTTCTGACCGAGGCTGCGCG 980
DB 705 TGGGCGCTTCTACACCCAGATTTGGGACCCCGGTGTGATCGGAGAGCGAGGCTGGG 764
QY 981 CCGGCTATGAGGCTTCCAGAGGCGCTGGGCGCTTCAAGCTCTTCAAGGAGGAACTT 1040
DB 765 GAGTGCCATAGATGCTTTTCAAGAGCGCGAGGGGTTTAAATTTACCTGAAAGACAT 824
QY 1041 CTCCATGCGCGGAGCCAGACATGAGGCTGCTGCTGCTGCGACTGGAGGAGCTCAT 1100
```

Search completed: July 13, 2004, 06:09:55
Job time : 6352.83 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 12, 2004, 20:25:00 ; Search time 4177.82 Seconds

(without alignments)

17647.909 Million cell updates/sec

Title: US-10-697-266-1

Perfect score: 2469

Sequence: 1 tgcggccaggtgtgcgg.....aaaaaaaaaaaaaaaaaaaaa 2469

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estrba.*

2: em_esthum.*

3: em_estmu.*

4: em_estmu.*

5: em_estov.*

6: em_estpl.*

7: em_estro.*

8: em_htc.*

9: gb_est1.*

10: gb_est2.*

11: gb_htc.*

12: gb_est3.*

13: gb_est4.*

14: gb_est5.*

15: em_estfun.*

16: em_estom.*

17: em_gss_hum.*

18: em_gss_inv.*

19: em_gss_pln.*

20: em_gss_vrt.*

21: em_gss_fun.*

22: em_gss_mam.*

23: em_gss_mus.*

24: em_gss_pro.*

25: em_gss_rod.*

26: em_gss_phg.*

27: em_gss_vrl.*

28: gb_gss1.*

29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	903.4	36.6	1070	13 BX400896	BX400896 BX400896
C 2	850	34.4	1017	13 BX444630	BX444630 BX444630
C 3	849.4	34.4	1201	13 BX447754	BX447754 BX447754
C 4	812.6	32.9	1201	9 AL578313	AL578313 AL578313

5	802.4	32.5	1001	13	BQ072325
6	796.4	32.3	1135	11	BC012135
7	787.6	31.9	899	13	BU171128
8	742	30.1	775	14	CA32432
9	722	29.2	1127	13	CA343432
10	716.6	29.0	1201	13	CB341691
11	715.8	29.0	1201	13	CB375208
12	665.2	26.9	703	14	CB851722
13	665	26.9	707	14	CB850862
14	626.2	25.4	650	13	CB850862
15	624.4	25.3	1201	13	CB850862
16	613.2	24.8	1243	12	CB850862
17	605.6	24.5	679	10	CB850862
18	588.8	23.8	607	14	CD513476
19	556	22.5	1067	12	BG339711
20	553.8	22.4	612	9	AI917311
21	544.4	22.0	3058	11	AK004849
22	541.2	21.9	3351	11	AK050214
23	533.6	21.6	830	12	BI518788
24	526	21.3	986	12	BI770906
25	519.2	21.0	611	10	BE328691
26	515.6	20.9	1109	12	BM548005
27	513.2	20.8	793	12	BI103148
28	508.6	20.6	542	10	AM874410
29	506	20.5	1201	9	AL555948
30	502.4	20.3	523	10	AM295415
31	502.4	20.3	926	13	BQ880416
32	477.8	19.4	493	12	BM193145
33	473.4	19.2	494	12	BM153139
34	468.6	19.0	709	10	BM408252
35	467.8	18.9	609	14	CB216237
36	466.6	18.9	506	9	AI624807
37	461	18.7	755	13	BQ179567
38	458.6	18.6	810	10	BE607519
39	446	18.1	1408	12	BM475275
40	437.6	17.7	490	10	BF348343
41	435.4	17.6	489	9	AA292035
42	426	17.3	1021	13	BX444631
43	423.2	17.1	721	13	BQ442874
44	411	16.6	955	13	BQ927119
45	403	16.3	999	12	BM802482

ALIGNMENTS

RESULT 1
BX400896/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BX400896 1070 bp mRNA linear EST 13-MAY-2003
BX400896 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
CDNA clone CS0DK004YE05 3-PRIME, mRNA sequence.

BX400896
CDNA clone CS0DK004YE05 3-PRIME, mRNA sequence.

BX400896.1 GI:30630225

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1070)

Li, W.B., Gruber, C., Jesse, J. and Polayes, D.

Full-length cDNA libraries and normalization

Unpublished (2001)

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of

Invitrogen. This sequence belongs to sequence cluster 4751.f For

more information about this cluster, see

http://www.genoscope.cns.fr/

cgi-bin/cluster.cgi?seq=CS0DK004AC03NP1&cluster=4751.f. Contact :

Feng Liang Email : fliang@lifetech.com URL :

http://fulllength.invitrogen.com/ Invitrogen Corporation 1600


```
Db 818 ASGGCTCACTGGCAAGTATGCGAGCTCGACCGTGAGGATGACTTCTGTGAGGCTGCCG 759
QY 1655 AGGCCCCGACATCCAGCCTAAGACCCACAGAGCCAGAGGCCAGATGCCACGCTGT 1714
Db 758 AGGCCCCGACATCCAGCCTAAGACCCACAGAGCCAGAGGCCAGATGCCACGCTGT 699
QY 1715 CCAGGGGAAGGGGCTGACATCTTCATCGCTGCGGCGCCCTGTGTGTCTCAGCCA 1774
Db 698 CCAGGGGAAGGGGCTGACATCTTCATCGCTGCGGCGCCCTGTGTGTCTCAGCCA 639
QY 1775 AGAACCGGTGGCGGCTGTGTGGGCGCGTCCACTGACCCAGAGAGGGGCGCTTTGGCC 1834
Db 638 AGAACCGGTGGCGGCTGTGTGGGCGCGTCCACTGACCCAGAGAGGGGCGCTTTGGCC 579
QY 1835 TCACGCTTTCGGGGAGACTCGCCTGTCTCATCGCTGCGCTCATCTTCAGGGAGCCAGGCGC 1894
Db 578 TCACGCTTTCGGGGAGACTCGCCTGTCTCATCGCTGCGCTCATCTTCAGGGAGCCAGGCGC 519
QY 1895 CGGCGGCTGGCTGAAGAGGGGCGACTACATTTGTGTGTCAGTGAATGGGAGCCATGAGGT 1954
Db 518 CGGCGGCTGGCTGAAGAGGGGCGACTACATTTGTGTGTCAGTGAATGGGAGCCATGAGGT 459
QY 1955 GGTGAGAGACACCGCAGAGGTGTGACGAGCTCAAGGCTCGCGGAGAGGGGCGCGCCAGCC 2014
Db 458 GGTGAGAGACACCGCAGAGGTGTGACGAGCTCAAGGCTCGCGGAGAGGGGCGCGCCAGCC 399
QY 2015 TGCAGGTGTGTGCTGTGCTGCCAGCTCTAGACTGCCCAAGCTTTGGGGAGCCCGCCCGC 2074
Db 398 TGCAGGTGTGTGCTGTGCTGCCAGCTCTAGACTGCCCAAGCTTTGGGGAGCCCGCCCGC 339
QY 2075 TCCTGCTGGGCCCCAGGGGGCTTTAAGAGCCAGAGGAGCATGTTGTCAAGACCCCGG 2134
Db 338 TCCTGCTGGGCCCCAGGGGGCTTTAAGAGCCAGAGGAGCATGTTGTCAAGACCCCGG 279
QY 2135 CATCCAGCTGGGCCAGTCCCGGCGCTCTCAACTGGAGCCGAAAGGCCAGCAGGGGCA 2194
Db 278 CATCCAGCTGGGCCAGTCCCGGCGCTCTCAACTGGAGCCGAAAGGCCAGCAGGGGCA 219
QY 2195 AGACTGGAGGTGCCCCCAGCCCTGTGCCCCAGTGAAGCAGCTCCGCTCATCTCTGA 2254
Db 218 AGACTGGAGGTGCCCCCAGCCCTGTGCCCCAGTGAAGCAGCTCCGCTCATCTCTGA 160
QY 2255 AGACCCAGAGGTGGCGCTGAGGCCAGATCTCGACCGCTCAGCCCTGAGCTCAGCT 2314
Db 159 AGACCCAGAGGTGGCGCTGAGGCCAGATCTCGACCGCTCAGCTCAGCTCAGCT 101
QY 2315 GGCAGCAAGCAGCAGCATGCCCTCCCGCAGAGGACCTCCGCGGCAATGCTGTCGG 2374
Db 100 GGCAGCAAGCAGCAGCATGCCCTCCCGCAGAGGAGCTCCCGGCAATGCTGTCGG 42
QY 2375 CCTCATGCTG 2384
Db 41 CCTCATGCTG 32
```

```
RESULT 3
BX447754/c
LOCUS
DEFINITION BX447754 Homo sapiens T CELLS (JURKAT CELL LINE) Homo sapiens cDNA
ACCESSION BX447754
VERSION BX447754.1 GI:31035860
KEYWORDS EST.
SOURCE Homo sapiens
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1201)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
```

Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
4751.f For more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0AH001CB05NP1&cluster=4751.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CS0AH001CB05NP1.

FEATURES source

1..1201
/location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DH001YD09"
/cell_line="JURKAT CELL LINE"
/tissue_type="T CELLS (JURKAT CELL LINE)"
/note="Vector: pcMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pcMVSPORT 6 vector.
Library was not normalized."

ORIGIN

Query Match 34.4%; Score 849.4; DB 13; Length 1201;
Best Local Similarity 95.6%; Pred. No. 5.9e-90;
Matches 936; Conservative 8; Mismatches 24; Indels 11; Gaps 7;
QY 1459 GAGGAGCGCAGCAGCTTGGCAAGCAGCACCTGAAAGCGTGCCATCTTGGGGCAGGAGG 1518
Db 976 GAAGAGCGCAGCAGCTTGGCAAG--MACACTTAGCGTGCCATCTTGGGGCAGGAGG 919
QY 1519 GCGCTGCGGCTGACGCGCC--TGTGCGCGCTCTGCGGAGGPGGACCTGCTGCGGCTGT 1577
Db 918 GCGCTGCGGCTGACGCGCCCTTGTGCGCGCTCTGCGGAGGTGACCTGCTGCGGCTGT 859
QY 1578 GATCTCCAGAGCGCTGACGCGCTCACTGCCAAGTATCGGAGCTCGACCTGAGGATGA 1637
Db 858 GATCTCCAGAG--GCTGCGAGCGCTCACTGGCCAAAGTATCGGAGCTCGACCGTGAGGATGA 800
QY 1638 CTTCTGTGAGGCTGCCAGGCGCCCGGACATCAGCCTTAAGACCCACAGAGCCAGAGGC 1697
Db 799 CTTCTGTGAGGCTGCCAGGCGCCCGGACATCAGCCTTAAGACCCACAGAGCCAGAGGC 740
QY 1698 CAGATGCGCAGCTGTCTCCAGGGGAAAGGGGCTGACATCTTCCATCGGCTGGGCGCCCT 1757
Db 739 CAGGATGCGCAGCTGTCTCCAGGGGAAAGGGGCTGACATCTTCCATCGGCTGGGCGCCCT 680
QY 1758 GTCTGTGTTCTCAGCCAGAACCGGTGGCGGCTGTGGGGCGCCGCTCCACCTGACCCGAG- 1816
Db 679 GTCTGTGTTCTCAGCCAGAACCGGTGGCGGCTGTGGGGCGCCGCTCCACCTGACCCGAG 620
QY 1817 GAGAGGGCGCTTTGGCCCTCAGCCTTGGGGAGAGCTGCGCTGTCTCATCGCTGCGGCTCA 1876
Db 619 GAGAGGGCGCTTTGGCCCTTACGCTTGGGGAGACTGCGCTGTCTCATCGCTGCGGCTCA 560
QY 1877 TTCAGGAGGCGCAGCGCGCGCTGCTGAAGAGGGGCGCACTACATTTGTGTCAGTGA 1936
Db 559 TTCAGGAGGCGCAGCGCGCGCTGCTGAAGAGGGGCGCACTACATTTGTGTCAGTGA 500
QY 1937 ATGGGCGAGCCATGAGGTGTGGAGACACCGGAGGTGGTGACGAGCTGAGGCTGCGG 1996
Db 499 ATGGGCGAGCCATGAGGTGTGGAGACACCGGAGGTGGTGACGAGCTGAGGCTGCGG 440
QY 1997 GAGAGGGGCGCGCAGCGCTGCGAGGTGTGTGCTGCTGCCAGCTCTAGACTGCGCCAGCT 2056
Db 439 GAGAGGGGCGCGCAGCGCTGCGAGGTGTGTGCTGCTGCCAGCTCTAGACTGCGCCAGCT 380
QY 2057 TGGGGGAGCGCGCGCGCTGCTGCTGCGGCGCCAGGGGCTTCTAAGAGGCGCAGAGGAGC 2116


```

VERSION      BQ072325.1  GI:19901371
KEYWORDS
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    NIH-MGC http://mgi.nci.nih.gov/.
AUTHORS      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
             Email: cgabbs-remail.nih.gov
             Tissue Procurement: Life Technologies, Inc.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
             found through the I.M.A.G.E. Consortium/LLNL at:
             http://image.llnl.gov
             Plate: LLAM12810 row: b column: 21
             High quality sequence stop: 623.
FEATURES     Location/Qualifiers
             1..1001
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /clone="IMAGE:5761460"
             /lab_host="DH10B"
             /clone_lib="NIH MGC 122"
             /note="organ: pooled lung and spleen; Vector: pCMV-SPORT6;
             Site 1: NotI; Site 2: EcoRV (destroyed); RNA source
             anonymous pool of 24 week female lung, 16 week female
             spleen, and 20-22 week male spleens. Library is oligo-dT
             primed and directionally cloned (EcoRV site is destroyed
             upon cloning). Average insert size 1.4 kb, insert size
             range 1-3 kb. Library is normalized and enriched for
             full-length clones and was constructed by C. Gruber
             (Invitrogen). Research Genetics tracking code 026. Note:
             this is a NIH_MGC Library."
ORIGIN
Query Match      32.5%; Score 802.4; DB 13; Length 1001;
Best Local Similarity 96.4%; Pred. No. 1.9e-84;
Matches 864; Conservative 0; Mismatches 26; Indels 6; Gaps 4;
QY 321 GGGCTGTGACTCCCTGACGACGATCCAGTGGCGCCAGCTGCAGAGCCGAGGCCACAGT 380
DB 42 GGGCTGTGACTCCCTGACGACGATCCAGTGGCGCCAGCTGCAGAGCCGAGGCCACAGT 101
QY 381 TCACCAGCAGATTGACAGGAGCTGCAGATGGGACGGCGCTGAGAACCTCTACAGAGC 440
DB 102 TCACCAGCAGATTGACAGGAGCTGCAGATGGGACGGCGCTGAGAACCTCTACAGAGC 161
QY 441 CACCAGCAACACCGGGTGCAGAGACGCTGCCTCGAGCTGAGTACGTCACCTCCAA 500
DB 162 CACCAGCAACACCGGGTGCAGAGACGCTGCCTCGAGCTGAGTACGTCACCTCCAA 221
QY 501 CCTGCAGCTGTGAAGGAGGAGCTGGAGAGCTCAGCGGTGGCGTGGACCTTGGCGGCA 560
DB 222 CCTGCAGCTGTGAAGGAGGAGCTGGAGAGCTCAGCGGTGGCGTGGACCTTGGCGGCA 281
QY 561 TGGGAGCGAAGCTGTACCTGTCCTCCATGATCCCTTGGCGCTGAGGAGACAAGAGGCT 620
DB 282 TGGGAGCGAAGCTGTACCTGTCCTCCATGATCCCTTGGCGCTGAGGAGACAAGAGGCT 341
QY 621 GGACTGGTCTACCGCTGAAGGAGCTGATCTCAGTGACATTTGGAGAGCGGGCTC 680
DB 342 GGACTGGTCTACCGCTGAAGGAGCTGATCTCAGTGACATTTGGAGAGCGGGCTC 401
QY 681 CTACGAGCAGAAATCAGGAGCTGGAGGCCCTTGGCGAGGCCATCGCGACCCCGAGCG 740
DB 402 CTACGAGCAGAAATCAGGAGCTGGAGGCCCTTGGCGAGGCCATCGCGACCCCGAGCG 461

```

```

QY 741 GAATGAGTCGGGCTGGAGCTGCTCACAGCTATTACAAACAGCTGTGCTTCTGTGATGC 800
DB 462 GAATGAGTCGGGCTGGAGCTGCTCACAGCTATTACAAACAGCTGTGCTTCTGTGATGC 521
QY 801 GGGCTTCTTCACCCCTGCCAGGAGCTCGGGCTCTTTTCCACTGTAAGACTCGCTTAC 860
DB 522 GGGCTTCTTCACCCCTGCCAGGAGCTCGGGCTCTTTTCCACTGTAAGACTCGCTTAC 581
QY 861 TGGGCTCCCGGCCAGCAGCGCTGCGCTTTCAGAGAGGCGAGCTTCTTCAACAT 920
DB 582 TGGGCTCCCGGCCAGCAGCGCTGCGCTTTCAGAGAGGCGAGCTTCTTCAACAT 641
QY 921 CGGTGCTTCCACACGACGAGATTGGGCGCGCCAGGACCGCTCTGACCGAGGGTGCCTG 980
DB 642 CGGTGCTTCCACACGACGAGATTGGGCGCGCCAGGACCGCTCTGACCGAGGGTGCCTG 701
QY 981 CGGCTGATGAGGCGCTTCCAGAGGCGCGCTGGGCGCTTACGCTTCTGAGGAGAACT 1039
DB 702 CGGCTGATGAGGCGCTTCCAGAGGCGCGCTGGGCGCTTACGCTTCTGAGGAGAACT 761
QY 1040 TCTCCATGCGCG-AGCCAGACATGAGCGCTGCGTCCCTCTGCGCACTGGAGAGCTC 1098
DB 762 TCTCCCTTGGCGGAGCCAGACATGAGCTGGTCCCTCTGGGCACTGGAGAGCTC 821
QY 1099 ATGATGGCCCGAGCCCGAGGAATGTGTGTTG--AGGGGCTCTTCAACACCTGCTCCATGG 1156
DB 822 ATGATGGCCCGAGCCCGAGGAATGTGGGGTTGGAAGGGCTCTTCAACACCTGCTCCATGG 881
QY 1157 CCCCCCAGACTGCTGG--CCAGCTGCGCTGGCGAGGAGCGGCCCGAGGCTG 1210
DB 882 GCCCCCCAGACTGCTGGGCTGGGCCCCAACCTGGCGCTTGGGCGAGGAAGCGCCCCGGGG 937

RESULT 6
BC012135
LOCUS      Homo sapiens rhophilin, Rho GTPase binding protein 1, mRNA (cdna
DEFINITION clone IMAGE:4555079), with apparent retained intron.
ACCESSION BC012135
VERSION    BC012135.1  GI:15082433
KEYWORDS   HTC.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 1135)
            Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
            Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
            Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
            Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
            Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
            Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
            Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
            Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
            Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
            McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
            Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
            Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
            Fahey, J., Helton, E., Kettman, M., Madan, A., Rodrigues, S.,
            Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
            Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
            Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
            Butlerfield, Y.S., Krzywinski, M.I., Skalska, U., Snailus, D.E.,
            Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
            Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
            Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
            22388257
            PUBMED 12477932
            2 (bases 1 to 1135)
            Strausberg, R.
            Direct Submission
            Submitted (02-AUG-2001) National Institutes of Health, Mammalian

```

Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: ATCC

CDNA Library Preparation: Rubin Laboratory

DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

BC Cancer Agency, Vancouver, BC, Canada

info@bcgsc.bc.ca

Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,

Susanna Chan, Readman Chiu, Chris Fiell, Erin Garland, Ran Guin,

Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo

Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven

Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline

Schein, Duane Smallus, Michael Smith, Lorraine Spence, Jeff Stott,

Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,

George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAL Plate: 29 Row: d Column: 17

This clone was selected for full length sequencing because it

passed the following selection criteria: matched mRNA gi: 19882236

This clone has the following problem: retained intron.

FEATURES

Location/Qualifiers
1..1135
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:455079"
/tissue_type="Uterus, leiomyosarcoma"
/clone_lib="NIH_MGC_46"
/lab_host="DH10B-R"
/note="Vector: pOTB7"

ORIGIN

Query Match 32.3%; Score 796.4; DB 11; Length 1135;
Best Local Similarity 99.9%; Pred. No. 9e-84;
Matches 797; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 321 GGGCTGTGACTCCCTGACGAGCTCCAGTGGCGGCGAGCTGCGAGCGCGAGGCCAGAT 380
DB 162 GGGCTGTGACTCCCTGACGAGATCCAGTGGCGGCGAGCTGCGAGCGCGAGGCCAGAT 221
QY 381 TCACACGAGATTGACAGGAGCTGCAGATGCGGACGGGCGCTGAGAACCTCTACAGAGC 440
DB 222 TCACACGAGATTGACAGGAGCTGCAGATGCGGACGGGCGCTGAGAACCTCTACAGAGC 281
QY 441 CACGAGCAACACCGGCTGAGAGACGGTGGCCCTGAGCTGAGCTACCTCACTCAA 500
DB 282 CACGAGCAACACCGGCTGAGAGACGGTGGCCCTGAGCTGAGCTACCTCACTCAA 341
QY 501 CTTGACCTGCTGAAGGAGAGCTGGAGGAGCTCAGCGGTGGCGGTGGACCTGGCGCGCA 560
DB 342 CTTGACCTGCTGAAGGAGAGCTGGAGGAGCTCAGCGGTGGCGGTGGACCTGGCGCGCA 401
QY 561 TGGAGCGAAGCTGTCTCTGTCCCATGATCCCTGGGCTGAGGAGACCAAGAGCT 620
DB 402 TGGAGCGAAGCTGTCTCTGTCCCATGATCCCTGGGCTGAGGAGACCAAGAGCT 461
QY 621 GGACTGCTTACACCGCTGAGGAGCTGATCTCAGTCACTTTTGGAGGAGCGCGCTC 680
DB 462 GGACTGCTTACACCGCTGAGGAGCTGATCTCAGTCACTTTTGGAGGAGCGCGCTC 521
QY 681 CTACGAGCGAGAAATCAGGAGCTGGAGCGCCCTCGGCGAGGCCCATCGGACCCCGAGCG 740
DB 522 CTACGAGCGAGAAATCAGGAGCTGGAGCGCCCTCGGCGAGGCCCATCGGACCCCGAGCG 581
QY 741 GAATGAGTCGGGCTGGAGCTGTCTCAGACCTATATACACGAGCTGTCTCTCTGGATGC 800

Db 582 GAATGAGTCGGGCTGGAGCTGTCTACAGCCTATTAAACACAGCTGTCTCTGGATGC 641
QY 801 GCGCTTCTCCACCCCTGCCAGAGCCTCGGGCTCTTCTTCCACTGTAGACTCGCTTAC 860
Db 642 GCGCTTCTCCACCCCTGCCAGAGCCTCGGGCTCTTCTTCCACTGTAGACTCGCTTAC 701
QY 861 TGGGGTCCCGGCGCCAGCAGCGTGGCCCTTCGAGAAAGGGCAGCGTTCTTTCACAT 920
Db 702 TGGGGTCCCGGCGCCAGCAGCGTGGCCCTTCGAGAAAGGGCAGCGTTCTTTCACAT 761
QY 921 CGGTGCGCTTCCACACGAGATTGGGGCGGCGCAGACCGGTCTCTGACCGAGGTCGCCG 980
Db 762 CGGTGCGCTTCCACACGAGATTGGGGCGGCGCAGACCGGTCTCTGACCGAGGTCGCCG 821
QY 981 CGCGCTATGGAGGCGCTTCCAGAGGCGCGTGGGGCTTTCAGGCTCTCTGAGGAGAACTT 1040
Db 822 CGCGCTATGGAGGCGCTTCCAGAGGCGCGTGGGGCTTTCAGGCTCTCTGAGGAGAACTT 881
QY 1041 CTCCCATGCGCGGCGCAGACATGAGCGTGGCTCCCTTTCGCGACTGGAGCAGCTCAT 1100
Db 882 CTCCCATGCGCGGCGCAGACATGAGCGTGGCTCCCTTTCGCGACTGGAGCAGCTCAT 941
QY 1101 GATGGCGCGGCGCCAGCA 1118
Db 942 GATGGCGCGGCGCCAGCA 959

RESULT 7

BU171128
LOCUS
DEFINITION AGENCOURT_7935660 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:614655
5', mRNA sequence.
ACCESSION BU171128
VERSION BU171128.1 GI:22685112
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 899)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>,
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM13475 Row: d column: 16
High quality sequence stop: 563.

FEATURES

source
1..899
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:614655"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_67"
/note="Organ: eye; Vector: pCMV-SPORT6; Site: 1; NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.75 kb. Library constructed by Life
Technologies."

ORIGIN

Query Match 31.9%; Score 787.6; DB 13; Length 899;
Best Local Similarity 98.6%; Pred. No. 1e-82;
Matches 826; Conservative 0; Mismatches 9; Indels 3; Gaps 3;

```
QY 1125 GTTTCAGGGCCCTTCACACACCTGCTCCATGCGCCGCCAAGACTGCTGGCCACGCTGGG 1184
Db 1 GTTTCAGGGCCCTTCACACACCTGCTCCATGCGCCGCCAAGACTGCTGGCCACGCTGGG 60

QY 1185 CTTGCGCAGGAGGGCCGCGCAGGTGCAGCCAGTACAGGCTAGTGCAACCGACCATGGC 1244
Db 61 CTTGCGCAGGAGGGCCGCGCAGGTGCAGCCAGTACAGGCTAGTGCAACCGACCATGGC 120

QY 1245 CGAGCCACCGTCCAGCACTAGTGCCTGTCTTCCTGGACTGCGCCCTGGTGCATGTCAAGGC 1304
Db 121 CGAGCCACCGTCCAGCACTAGTGCCTGTCTTCCTGGACTGCGCCCTGGTGCATGTCAAGGC 180

QY 1305 CGAGTACTTCGCTCCCTGGCCCACTACACATGACCATGGCCCTTCGGACGGCTCCCC 1364
Db 181 CGAGTACTTCGCTCCCTGGCCCACTACACATGACCATGGCCCTTCGGACGGCTCCCC 240

QY 1365 AGCGACCGAGGAGAGCTCCCCACGACGAGCAGTCTTCCTGACGCCCCCACCCTCTCTC 1424
Db 241 AGCGACCGAGGAGAGCTCCCCACGACGAGCAGTCTTCCTGACGCCCCCACCCTCTCTC 300

QY 1425 TAAGCCCGAGGCCCTGTGTGTCGCGCAGAGCTGAGGAGGCGCAGGAGCTTGGCAAGGC 1484
Db 301 TAAGCCCGAGGCCCTGTGTGTCGCGCAGAGCTGAGGAGGCGCAGGAGCTTGGCAAGGC 360

QY 1485 ACACCTGAAGGTGCTCATCTCGGGCAGGAGGCGCTGGGCTGCGAGCCCTGTGCGG 1544
Db 361 ACACCTGAAGGTGCTCATCTCGGGCAGGAGGCGCTGGGCTGCGAGCCCTGTGCGG 420

QY 1545 CGTCTGCGAGGTGAGCTGCTTCGCGGCTGTGATCTCCAGAGCGCTGCAGCGCTCACT 1604
Db 421 CGTCTGCGAGGTGAGCTGCTTCGCGGCTGTGATCTCCAGAGCGCTGCAGCGCTCACT 480

QY 1605 GCGCAAGTATGCGAGCTGCGAGTGCAGCTGAGTGAATCTTGTGAGGCTGCGAGGCCCGGA 1664
Db 481 GCGCAAGTATGCGAGCTGCGAGTGCAGCTGAGTGAATCTTGTGAGGCTGCGAGGCCCGGA 540

QY 1665 CATCAGCCTAAGACCCACAGAGCCAGAGCCAGGATGCCAGCCTGTCCAGGGGAA 1724
Db 541 CATCAGCCTAAGACCCACAGAGCCAGAGCCAGGATGCCAGCCTGTCCAGGGGAA 600

QY 1725 GGGG-CCTGACATCTTCCATCGGCTGGGGCCCTGTCTGTGTCTCAGCCAAAGAACCGGT 1783
Db 601 GGGGCCCTGACATCTTCCATCGGCTGGGGCCCTGTCTGTGTCTCAGCCAAAGAACCGGT 660

QY 1784 GCGCGCTGGTGGGCGCCCTGCACCTGACCCGAGGAGAGGGCGGCTTTGGCTTCACGCTTC 1843
Db 661 GCGCGCTGGTGGGCGCCCTGCACCTGACCCGAGGAGAGGGCGGCTTTGGCTTCACGCTTC 720

QY 1844 GGGGAGACTCGCCTGTCTCTCATGCTGCGCTCATTTCCAGGAGCCAGGCGCGG-CGGCT 1902
Db 721 GGGGAGACTCGCCTGTCTCTCATGCTGCGCTCATTTCCAGGAGCCAGGCGCGCGGTT 780

QY 1903 GGCCTGAAGGAGGGCGACTACATT-GTGTCAGTGAATGGGCGAGCCATCAGGTGGTGG 1959
Db 781 GGCCTGAAGGAGGGCGACTACATTGTTGGGTCAGTGAATGGGCGAGCCATGCGAGGGGGG 838

RESULT 8
CA432432/c
LOCUS 775 bp mRNA linear EST 08-NOV-2002
DEFINITION UI-H-FL1-bgu-e-15-0-UI-s1 NCI CGAP_Fl1 Homo sapiens CDNA clone
            UI-H-FL1-bgu-e-15-0-UI 3', mRNA sequence.
ACCESSION CA432432
VERSION CA432432.1 GI:24795158
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS 1 (bases 1 to 775)
          NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
          Tumor Gene Index
```

JOURNAL
COMMENT

Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: James Martin
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@uiowa.edu
Seq primer: M13 FORWARD
POLYA=Yes

FEATURES
source

Location/Qualifiers
1..775
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-FL1-bgu-e-15-0-UI"
/tissue_type="Cell lines"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP_Fl1"
/note=Torgan: Chondrosarcoma; Vector: pT7T3-Pac
(Pharmacia) with a modified polylinker; Site 1: EcoR I;
Site 2: Not I; NCI CGAP_Fl1 is a normalized cDNA library
derived from a pool of mRNA obtained from 4 cell lines
from grade III chondrosarcoma tissues. The library was
constructed according to Bonaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an EcoR I
adaptor, digested with Not I, and cloned directionally
into pT7T3-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dT)18 tail. The sequence tag for this library is
GAGTCGGTG. The cell lines were provided by Dr. James
Martin from the University of Iowa.
TAG TISSUE=Human Chondrosarcoma Grade 3 cell line mix
TAG LIB=UI-H-FL1
TAG_SEQ=GAGTCGGTG"

ORIGIN

Query Match 30.1%; Score 742; DB 14; Length 775;
Best Local Similarity 99.6%; Pred. No. 2.2e-77;
Matches 775; Conservative 0; Mismatches 0; Indels 3; Gaps 3;

QY 1677 GACCCACAGAGCCAGAGCCAGGATGCCAGCTGTGCCAGGGAGGGGCTGACAT 1736
Db 775 GACCCACAGAGCCAGAGCCAGGATGCCAGCTGTGCCAGGGAGGGGCTGACAT 716

QY 1737 CTTCCATCGGCTGGGGCCCTCTGTCTGTCTCAGCCAAAGAACCGGTGGCGCTGGTGGG 1796
Db 715 CTTCCATCGGCT-GGGCCCTGTCTGT-TTCTCAGCCAGAACCGGTGGCGCTGGTGGG 658

QY 1797 GCGCGTCCACTGACCCGAGAGAGGGCGGCTTTGGGCTCAGCTTCGGGGAGACTCGCC 1856
Db 657 GCGCGTCCACTGACCCGAGAGAGGGCGGCTTTGGGCTCAGCTTCGGGGAGACTCGCC 598

QY 1857 TGTCTCATCGCTCGCTCATTTCCAGGAGCCAGGCGCGGCTGGCTTGAAGGAGGG 1916
Db 597 TGTCTCATCGCTCGCTCATTTCCAGGAGCCAGGCGCGGCTGGCTTGAAGGAGGG 538

QY 1917 CGACTACATTTGTTCAGTGAATGGGCGAGCCATGCAAGTGGTGGAGACACGCGAGGTGGT 1976
Db 537 CGACTACATTTGTTCAGTGAATGGGCGAGCCATGCAAGTGGTGGAGACACGCGAGGTGGT 478

QY 1977 GACGAGCTGAAGGCTCGGGAGAGCGCGCCAGCTTCAGGTGGTGGTGGTGGTGGC 2036
Db 477 GACGAGCTGAAGGCTCGGGAGAGCGCGCGCCAGCTTCAGGTGGTGGTGGTGGTGGC 418

QY 2037 CAGCTCTAGACTGCCAGCTTGGGGAGCCGCGCGCGCTCTCTGCTGGGCCCCCAGGGGGCT 2096
Db 417 CAGCTCTAGACTGCCAGCTTGGGGAGCCGCGCGCGCTCTCTGCTGGGCCCCCAGGGGGCT 358

```
QY 2097 TCTAAGGACGAGGGAGCATGGTTGCAAGACCCCGGCATCCAGTGGGCGAGTCCCCG 2156
Db 357 TCTAAGGAGCCAGAGGAGCATGGTTGCAAGACCCCGGCATCCAGTGGGCGAGTCCCCG 298
QY 2157 GCCCTCTCTCAACTGAGCGGAAAGCCAGCAGGCGCAAGACTGGAGGCTGCCCGCCAGCC 2216
Db 297 GCCCTCTCTCAACTGAGCGGAAAGCCAGCAGGCGCAAGACTGGAGGCTGCCCGCCAGCC 238
QY 2217 CTGTGCCCCAGTGAAGCAGCTCCGGCCTCATCTTTGAAGCACCCAGGGTGGCGGTGAGG 2276
Db 237 CTGTGCCCCAGTGAAGCAGCTCCGGCCTCATCTTTGAAGCACCCAGGGTGGCGGTGAGG 178
QY 2277 GCCAGATCCCTGACGCCCTGAGCCTGCTCCAGCTGCGCAGCAAGCACCGAGCATGCC 2336
Db 177 GCCAGATCCCTGACG - CCTCAGCCCTGGCTCCAGCTGCGCAGCAAGCACCGAGCATGCC 119
QY 2337 CTCCCCACCCAGAGGACCTCCGGGCAATGCTGTCCGCCCTCATGCTGGAGGCTGCCCTCG 2396
Db 118 CTCCCCACCCAGAGGACCTCCGGGCAATGCTGTCCGCCCTCATGCTGGAGGCTGCCCTCG 59
QY 2397 GGCACCTGCTGCCATTAAGACTGTGTGAGACTGTCTGAAAAAATAAAAAA 2454
Db 58 GGCACCTGCTGCCATTAAGACTGTGTGAGACTGTCTGAAAAAATAAAAAA 1
```

```
RESULT 9
BX341691 1127 bp mRNA linear EST 01-MAY-2003
LOCUS BX341691 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
DEFINITION cDNA clone CS0DK004YE05 5-PRIME, mRNA sequence.
ACCESSION BX341691
VERSION BX341691
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1127)
AUTHORS Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 4751.f For
more information about this cluster, see
http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0DK004AC03QF1
cluster=4751.f Contact : feng liang Email : fliang@lifetech.com
URL : http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DK004AC03QF1.
```

```
FEATURES
source
1..1127
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DK004YE05"
/cell_type="HELA CELLS COT 25-NORMALIZED"
/cell_line="HELA"
/clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
```

ORIGIN

```
Query Match 29.2%; Score 722; DB 13; Length 1127;
Best Local Similarity 95.5%; Pred. No. 4.2e-75;
Matches 806; Conservative 10; Mismatches 20; Indels 8; Gaps 7;
QY 321 GGGCTGTGACTCCCTGACGAGATCCAGTGGCGCCAGCTGCAGAGCCGAGGCCAGAT 380
```

```
Db 267 GGGCTGTGACTCCCTGACGAGATCCAGTGGCGCCAGCTGCAGAGCCGCA -GGCCCCAGAT 325
QY 381 TCACACGAGATGACAAAGAGCTGCAGATGCGGACGGGGCTGAGAACTCTTACAGAGC 440
Db 326 TCACACGAGATGACAAAGAGCTGCAGATGCGGACGGGGCTGAGAACTCTTACAGAGC 385
QY 441 CACCACAAACACCGGGTGCAGAGAGCGGTGCGCCCTGGAGCTGAGCTACGTCACATCCAA 500
Db 386 CACCACAAACACCGGGTGCAGAGAGCGGTGCGCCCTGGAGCTGAGCTACGTCACATCCAA 445
QY 501 CTTGAGCTGCTGAAAGAGAGCTGCAGAGAGCTCAGCGGTGGCGTGGACCTGGCGGCA 560
Db 446 CTTGAGCTGCTGAAAGAGAGCTGCAGAGAGCTCAGCGGTGGCGTGGACCTGGCGGCA 505
QY 561 TGGGAGGAGAGCTGCTCACTGTCCCATGATCCCGCTGGGCTGAAGGAGACCAAGAGCT 620
Db 506 TGGGAGGAGAA -CTGTCACTGTCCCATGATCCCGCTGGGCTGAAGGAGACCAAGAGCT 564
QY 621 GGAAGTGTGTACACCCCTGAAAGAGAGCTGATCTCAGTGCACCTTTGGAGAGAGCGGCTC 680
Db 565 GGAAGTGTGTACACCCCTGAAAGAGAGCTGATCTCAGTGCACCTTTGGAGAGAGCGGCTC 624
QY 681 CTACGAGGAGAGAAATCAGGAGAGCTGAGGAGCGCTGCGGAGGCCATCGGAGACCCAGCGG 740
Db 625 CTACGAGGAGAGAAATCAGGAGAGCTGAGGAGCGCTGCGGAGGCCATCGGAGACCCAGCGG 684
QY 741 GAATGAGTGGGCGCTGGAGCTGCTCAGAGCTATTACAAACAGCTGTGCTTCTGATGTC 800
Db 685 GAATGAGTGGGCGCTGGAGCTGCTCAGAGCTATTACAAACAGCTGTGCTTCTGATGTC 744
QY 801 GCGCTTCTCACCACCCCTGCCAGAGAGCTCGGGCTCTTCTTCCACTGGTACGACTCGCTTAC 860
Db 745 GCGCTTCTCACCACCCCTGCCAGAGAGCTCGGGCTCTTCTTCCACTGGTACGACTCGCTTAC 804
QY 861 TGGGGTCCCGGCCACAGAGCGTCCCTGGCTTCGAGAGAGGAGCGGCTTCTTCAACAT 920
Db 805 TGGGGT -CCGGGCCACAGCGTCCCTGGCTTCGAGAGAGGAGCGGCTTCTTCAACAT 863
QY 921 CGGTGCCCTCCACACGACAGATTGGGGCGGCCAGGACCGCTCTCGACACGAGGCTGCCCG 980
Db 864 CGGTGCCCTCCACACGACAGATTGGGGCGGC -CAGGACCGCTCTCTGCCAGAGGT -SCCG 921
QY 981 CCGCGCTATGGAGGCGCTTCCAGAGGGCGCTGGGGCGCTTCAGCCTCTGAGGGAGAACTT 1040
Db 922 CCGCGCTATGGAGG -CTTCCAGAGGGCGCTGGGGCGCTTCAGCCTCTGAGGGAGAACTT 980
QY 1041 CTCCTATGCGCCGAGCGCCAGACATGAGCGCTGCGTCCCTCTGCGCACTGGAGAGCTCAT 1100
Db 981 CTCCTATGCGCCGAGCGCCAGACATGAGCGCTGCGTCCCTCTGCGCACTGGAGAGAGCTCAT 1040
QY 1101 GATGGCCCGAGCCCGAGAAATGCTGTTTGGAGGCGCTCTCACCACCTGCTCCATGGGCCCC 1160
Db 1041 GATGGCCCGAGCCCGAGAAATGCTGTTTGGAGGCGCTCTCACCACCTGCTCCATGGGCCCC 1098
QY 1161 CCAA 1164
Db 1099 CAAA 1102
```

```
RESULT 10
BX375208 1201 bp mRNA linear EST 08-MAY-2003
LOCUS BX375208 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
DEFINITION cDNA clone CS0DC010YG11 5-PRIME, mRNA sequence.
ACCESSION BX375208
VERSION BX375208
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1201)
```

```

AUTHORS      Li.W.B., Gruber,C., Jesse,J. and Polayes,D.
TITLE        Full-length cDNA libraries and normalization
JOURNAL      Unpublished (2001)
COMMENT      Contact: Genoscope
              Genoscope - Centre National de Sequencage
              BP 191 91006 EVRY cedex - France
              Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
              Library was constructed by Life Technologies, a division of
              Invitrogen. This sequence belongs to sequence cluster 4751.f For
              more information about this cluster, see
              http://www.genoscope.cns.fr/
              cgi-bin/cluster.cgi?seq=CS0DH001AD06QP1&cluster=4751.f. Contact :
              Feng Liang Email : fliang@lifetech.com URL :
              http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
              Faraday Avenue Genoscope sequence ID : CS0DH001AD06QP1.
FEATURES     Location/Qualifiers
              1..1201
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone="CS0DH001YD09"
              /tissue_type="T CELLS (JURKAT CELL LINE)"
              /cell_line="JURKAT CELL LINE"
              /clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE)"
              /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
              with a NotI-oligo (dT) primer. Five prime end enriched,
              double-strand cDNA was digested with Not I and cloned into
              the Not I and EcoRV sites of the pCMVSPORT 6 vector.
              Library was not normalized."
ORIGIN
source
Query Match      29.0%; Score 716.6; DB 13; Length 1201;
Best Local Similarity 94.5%; Pred. No. 1.7e-74;
Matches 778; Conservative 16; Mismatches 20; Indels 9; Gaps 5;
QY 321 GGGCTGTGACTCCCTGACGACATCAGTGGCGCCAGCTGCAGAGCCGAGGCCAGAT 380
DB 223 GGGCTGTGACTCCCTGACGACATCAGTGGCGCCAGCTGCAGAGCCGAGGCCAGAT 282
QY 381 TCACGACAGATTGACAGGAGCTCAGATGGGAGCGGCGCTGAGAACCTTACAGAGC 440
DB 283 TCACGACAGATTGACAGGAGCTCAGATGGGAGCGGCGCTGAGAACCTTACAGAGC 342
QY 441 CACGAGCAACACCGGTGAGAGACGCTGCCCTGGAGCTGAGTCACTCACTCCAA 500
DB 343 CACGAGCAACACCGGTGAGAGACGCTGCCCTGGAGCTGAGTCACTCACTCCAA 402
QY 501 CTGCAGCTGCTGAAGGAGGAGCTGAGGAGCTCAGCGGTGGCGTGGACCTTGGCCGGCA 560
DB 403 CTGCAGCTGCTGAAGGAGGAGCTGAGGAGCTCAGCGGTGGCGTGGACCTTGGCCGGCA 462
QY 561 TGGAGCGAAGCTGTACTTCCCATGATCCCTTGGGCTGGAAGGAGCAAGAGGCT 620
DB 463 TGGAGCGAAG-CTGTACTTCCCATGATCCCTTGGGCTGGAAGGAGCAAGAGGCT 521
QY 621 GGACTGTCTCACCGCTGAAGGAGCTGATCTCAGTGACCTTTGAGAGGAGCGGCGCTC 680
DB 522 GGACTGTCTCACCGCTGAAGGAGCTGATCTCAGTGACCTTTGAGAGGAGCGGCGCTC 581
QY 681 CTACGAGCAGAAATCAGGAGCTGGAGCCCTGGGAGGAGCCATCGAGACCCCGAGCGG 740
DB 582 CTACGAGCAGAAATCAGGAGCTGGAGCCCTGGGAGGAGCCATCGAGACCCCGAGCGG 641
QY 741 GAATGAGTCGGGCTGGAGCTGCTCAGAGCTATTATCAACAGCTGTGCTTCTGGATGC 800
DB 642 GAATGAGTCGGGCTGGAGCTGCTCAGAGCTATTATCAACAGCTGTGCTTCTGGATGC 701
QY 801 GCGCTTCTCAGCCCTGACAGAGCTCGGGCTCTTCTTCCACTGCTAGCTGCTTAC 860
DB 702 GCGCTTCTCAGCCCTGACAGAGCTCGGGCTCTTCTTCCACTGCTAGCTGCTTAC 761
QY 861 TGGGCTCCCGGCCACAGCAGCTGCGCTTTCGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 920
DB 762 TGGGCTCCCGGCCACAGCAGCTGCGCTTTCGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 821

```

```

QY 921 CGGTGCTCCACACGACGAGA-TTGGGGCGCGCCAGAGACCGCTCTCTGACACGAGGTGCC 979
DB 822 CGGTGCTCCACACGACGAGA-TTGGGGCGCGCCAGAGACCGCTCTCTGACACGAGGTGCC 881
QY 980 GCGGCGCTATGAGAGCGCTTCCAGAGGGCGCTGGGGCCTTACGCTCTCTGAGGAGAACT 1039
DB 882 GCGGCGCTATGAGAGCGCTTCCAGAGGGCGCTGGGGCCTTACGCTCTCTGAGGAGAA--C 939
QY 1040 TCTCCCATGCGCGGAGCCAGACATGAGCGCTGCGTCCCTCTGCGCACTGGAGAGCTCA 1099
DB 940 TTYTCAATGCTSCGAGCCARACATGAGCGCTS---STCCCTCTGCCACTGGAGAGCTCA 996
QY 1100 TGATGGCCAGGCCCCAGGAATGTGTGTTGAGGGCTCTCACC 1142
DB 997 TGATGGCCW--GGCCARGAATGTTKTTTTRAGGCTTTCACMC 1037

RESULT 11
BX405203      1201 bp      mRNA      linear      EST 13-MAY-2003
DEFINITION   clone CS0DH001YD09 5-PRIME, mRNA sequence.
ACCESSION    BX405203
VERSION      BX405203.1 GI:30635313
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1 (bases 1 to 1201)
AUTHORS      Li.W.B., Gruber,C., Jesse,J. and Polayes,D.
TITLE        Full-length cDNA libraries and normalization
JOURNAL      Unpublished (2001)
COMMENT      Contact: Genoscope
              Genoscope - Centre National de Sequencage
              BP 191 91006 EVRY cedex - France
              Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
              Library was constructed by Life Technologies, a division of
              Invitrogen. This sequence belongs to sequence cluster 4751.f For
              more information about this cluster, see
              http://www.genoscope.cns.fr/
              cgi-bin/cluster.cgi?seq=CS0AH001CB05QP1&cluster=4751.f. Contact :
              Feng Liang Email : fliang@lifetech.com URL :
              http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
              Faraday Avenue Genoscope sequence ID : CS0AH001CB05QP1.
FEATURES     Location/Qualifiers
              1..1201
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone="CS0DH001YD09"
              /tissue_type="T CELLS (JURKAT CELL LINE)"
              /cell_line="JURKAT CELL LINE"
              /clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE)"
              /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
              with a NotI-oligo (dT) primer. Five prime end enriched,
              double-strand cDNA was digested with Not I and cloned into
              the Not I and EcoRV sites of the pCMVSPORT 6 vector.
              Library was not normalized."
ORIGIN
Query Match      29.0%; Score 715.8; DB 13; Length 1201;
Best Local Similarity 91.9%; Pred. No. 2.1e-74;
Matches 824; Conservative 6; Mismatches 13; Indels 54; Gaps 5;
QY 329 ACTCCTGACCCAGATCCAGTGGCGCCAGCTGCAGAGCCGAGGCCAGATTACACGAC 388
DB 211 ACCCGCGCTCAGATCCAGTGGCGCCAGCTGCAGAGCCGCA-GGCCAGATTACACGAC 269
QY 389 AGATTGACAAGGAGCTGCAGATGGAGCGGCGCTGAGAACTCTACAGAGCCACGACGA 448
DB 270 AGATTGACAAGGAGCTGCAGATGGAGCGGCGCTGAGAACTCTACAGAGCCACGACGA 329

```

```
QY 449 ACAACCGGTGAGAGAGCGTCCCTGGAGCTGAGCTACGTCAACTCAACCTGCAGC 508
Db 330 ACAACCGGTGAGAGAGCGTCCCTGGAGCTGAGCTACGTCAACTCAACCTGCAGC 389
QY 509 TGTCTGAAGGAGGAGCTGGAGAGCTCAGCGGTGGCGTGGACCCCTGGCCGCGCATGGGAGCG 568
Db 390 TGTCTGAAGGAGGAGCTGGAGAGCTCAGCGGTGGCGTGGACCCCTGGCCGCGCATGGGAGCG 449
QY 569 AAGCTGTACTGTCCCATGATCCCTTGGGCTTGAAGAGAGACAGAGAGCTGGACTGGT 628
Db 450 AA-CTGTCACTGTCCCATGATCCCTTGGGCTTGAAGAGAGACTAAGAGAGCTGGACTGGT 508
QY 629 CTACACCGCTGAAGGAGCTGATCTCAGTGACATTTGGAGAGGAGCGCGCTCTACGAGG 688
Db 509 CTACACCGCTGAAGGAGCTGATCTCAGTGACATTTGGAGAGGAGCGCGCTCTACGAGG 568
QY 689 CAGAAATCAGGAGAGCTGGAGGCGCTCGGCGAGGCCATTCGAGACCCCGCAGCGGAATGAGT 748
Db 569 CAGAAATCAGGAGAGCTGGAGGCGCTCGGCGAGGCCATTCGAGACCCCGCAGCGGAATGAGT 628
QY 749 CGGGCTTGGAGCTGCTCAGACGCTATTAACACAGCTGTCTTCTCTGGATGGCGGCTTCC 808
Db 629 CGGGCTTGGAGCTGCTCAGACGCTATTAACACAGCTGTCTTCTCTGGATGGCGGCTTCC 688
QY 809 TCACCCCTGCGAGAGGCTCGGCGCTCTTCTTCACTGGTACGACTCGCTTACTGGGCTCC 868
Db 689 TCACCCCTGCGAGAGGCTCGGCGCTCTTCTTCACTGGTACGACTCGCTTACTGGGCTCC 748
QY 869 CGGCCCGCAGAGCTGCGCTTGGGCTTGGAGAGGAGCGCTTCTTCAACATCGGTGCC 928
Db 749 CGGCCCGCAGAGCTGCGCTTGGGCTTGGAGAGGAGCGCTTCTTCAACATCGGTGCC 808
QY 929 TCACACCGCAGATTTGGGGCGCGCAGGACCGCTCTTGCACCGAGGAGTGGCGCGCGCT 987
Db 809 TCACACCGCAGATTTGGGGCGCGCAGGACCGCTCTTGCACCGAGGAGTGGCGCGCGCT 868
QY 988 ATGAGGCGCTTCAGAGGCGCGCTGGGGCGCTTCAGCGCTCTGAGGGAGAACTTCTCCCAT 1047
Db 869 ATGAGGCGCTTCAGAGGCGCGCT----- 892
QY 1048 GCGCCGAGCCAGACATAGCGCTGCGTCCCTCTGGCCTTGGAGCAGCTCATGATGCC 1107
Db 893 -----GACATAGCGCTGCGTCCCTCTGCGCACTGGAGCAGCTCATGATGCC 940
QY 1108 CAGGCCCGAATGTGTTTGGGGCTCTCACCACCTGCTCCATGCGCCCGCCAGAC 1167
Db 941 CAGGCCCGAATGTGTTTGGGGCTCTCACCACCTGCTCCATGCGCCCGCCAGAC 1000
QY 1168 TGCTGGCCCGAGCTGCGCTTGGGGCGCGCAGGAGCGCGCCAGGTGGCAGCGGAGTACAGG 1224
Db 1001 TGCTGGCCCGAGCTGCGC---TGGSGRGAAGCGCGCCARGTGGCAGCGGAGTACAGS 1054

RESULT 12
CB851722/c
LOCUS
DEFINITION
UI-CF-EN1-aeK-a-17-0-UI-el UI-CF-EN1 Homo sapiens cDNA clone
UI-CF-EN1-aeK-a-17-0-UI-3', mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS
TITLE
Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL
Genome Res. 6 (9), 791-806 (1996)
MEDLINE
97044477
PUBMED
8889548
COMMENT
Contact: McCray, PB
```

```
McCray Lab
University of Iowa
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
Tel: 319 356 4866
Fax: 319 356 7171
Email: paul-mccray@uiowa.edu
Tissue Procurement: Dr. M. J. Welsh, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com) or from Open Biosystems
(www.openbiosystems.com)
Seq primer: M13 FORWARD
POLYA=Yes
Location/Qualifiers
1. .703
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-CF-EN1-aeK-a-17-0-UI"
/tissue_type="Primary Lung Cystic Fibrosis Epithelial
Cells"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-CF-EN1"
/notes="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a
modified polylinker; Site_1: EcoR I; Site_2: Not I;
UI-CF-EN1 is a normalized cDNA library containing the
following tissue(s): Primary Lung Cystic Fibrosis
Epithelial Cells. The library was constructed according to
Bonaldi, Lennon and Soares, Genome Research, 6:791-806,
1996. First strand cDNA synthesis was primed with an
oligo-dT primer containing a Not I site. Double stranded
cDNA was ligated to an EcoR I adaptor, digested with Not
I, and cloned directionally into pT73-Pac vector. The
oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tag for this library is CTGCTCAGGT.
TAG_SEQ=None found"
```

ORIGIN

```
Query Match 26.9%; Score 665.2; DB 14; Length 703;
Best Local Similarity 99.2%; Pred. No. 2e-68;
Matches 700; Conservative 0; Mismatches 3; Indels 3; Gaps 3;
QY 1747 CTGGGGCCCTGTCTGTGTTCTCAGGCAAGAACCGGTGGCGCTGTGGGGCCGCTCCAC 1806
Db 703 CAGGGGCCCCGTGTCTGTG-TCTCAGCCAGAACCGGTGGCGCTGTG-TGGGGCCGCTCCAC 646
QY 1807 CTGACCCGAGGAGGCGCGCTTTGGCTCAGCTTCGGGGAGACTCGCTGTCTCTCATC 1866
Db 645 CTGACCCGAGGAGGCGCGCTTTGGCTCAGCTTCGGGGAGACTCGCTGTCTCTCATC 586
QY 1867 GCTGCCCGCTCATTCAGGAGCCAGGCGCGCGCTGGCTTGAAGGAGGCGCATACATT 1926
Db 585 GCTGCCCGCTCATTCAGGAGCCAGGCGCGCGCTGGCTTGAAGGAGGCGCATACATT 526
QY 1927 GTGTCAAGTGAATGGGAGCCATCAGGTGTGTGAGACACCGGAGGTGGTGAAGGAGCTG 1986
Db 525 GTGTCAAGTGAATGGGAGCCATCAGGTGTGTGAGACACCGGAGGTGGTGAAGGAGCTG 466
QY 1987 AAGGCTTCGGGAGAGGCGCGCGCTTCAGCTTCAGTGTGTGCTGTGCTGCCAGCTCTAGA 2046
Db 465 AAGGCTTCGGGAGAGGCGCGCGCTTCAGCTTCAGTGTGTGCTGTGCTGCCAGCTCTAGA 406
QY 2047 CTGCCAGCTTGGGGAGACCGCGCGCGCTTCTGTCTGGGCGCCCGAGGGGCTTTAAGGAGC 2106
Db 405 CTGCCAGCTTGGGGAGACCGCGCGCGCTTCTGTCTGGGCGCCCGAGGGGCTTTAAGGAGC 346
QY 2107 CAGAGGAGGAGCATGTTGCRAGAGCCCGGAGCTCCAGCTGGGCGGAGTCCCGGCGCTCTC 2166
|||||
```


SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 650)
 AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaps-k@mail.nih.gov
 Tissue Procurement: James Martin
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Clone distribution information can be obtained from Dr. M. Bento Soares, bento-soares@uiowa.edu
 Seq Primer: M13 FORWARD
 POLYA=Yes

FEATURES
 source Location/Qualifiers
 1..650
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="UI-H-FLO-bdr-a-08-0-UI"
 /tissue_type="Cell lines"
 /dev_stage="Adult"
 /lab_host="DH10B (Life Technologies)"
 /clone_lib="NCI CGAP FLO"
 /note="Organ: Chondrosarcoma; Vector: p7T3-Pac (Pharmacia) with a modified polylinker; Site 1: Ecor I; Site 2: Not I; NCI CGAP FLO is a cDNA library derived from a pool of mRNA obtained from 4 cell lines from grade III chondrosarcoma tissues. The library was constructed according to Bonaldo, Lemmon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an Ecor I adaptor, digested with Not I, and cloned directionally into p7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is GAGTCCGTTG. The cell line was provided by Dr. James Martin from University of Iowa.
 TAG_TISSUE=Human Chondrosarcoma grade 3 cell line mix
 TAG_LIB=UI-H-FLO
 TAG_SEQ=GAGTCCGTTG"

ORIGIN
 Query Match 25.4%; Score 626.2; DB 13; Length 650;
 Best Local Similarity 98.6%; Pred. No. 7.2e-64;
 Matches 642; Conservative 0; Mismatches 8; Indels 1; Gaps 1;
 QY 1809 GACCCGAGGAGGCGGCTTTGGCCTCAGCTTCGGGGAGACTCGCCTGTCTCATCGC 1868
 Db 650 GACCCGAGGAGGCGGCTTTGGCCTCAGCTTCGGGGAGACTCGCCTGTCTCATCGC 591
 QY 1869 TGCCGTATTTACGGAGCCACGAGCCGCGGCTGTGGCTGAAGAGGCGGACTACATGT 1928
 Db 590 TGCCGTATTTACGGAGCCACGAGCCGCGGCTGTGGCTGAAGAGGCGGACTACATGT 531
 QY 1929 GTCACTGAATGGCAGCATTCAGTGTGTGAGACACGCGAGTGTGACGGAGCTGAA 1988
 Db 530 GTCACTGAATGGCAGCATTCAGTGTGTGAGACACGCGAGTGTGACGGAGCTGAA 471
 QY 1989 GGCTCGGGAGAGGCGGCGCCAGCTTCAGTGTGTGCTGTGCCAGCTTAGACT 2048
 Db 470 GGCTCGGGAGAGGCGGCGCCAGCTTCAGTGTGTGCTGTGCCAGCTTAGACT 411
 QY 2049 GCCAGCTTGGGGNACCGCGGCGCTTCCTCGGGCCCGAGGGGCTTCTAAGAGGCA 2108
 Db 410 GCCAGCTTGGGGNACCGCGGCGCTTCCTCGGGCCCGAGGGGCTTCTAAGAGGCA 351

QY 2109 GAGGGAGCATGTTGCAAGACCCCGGCATCCAGTGGGGCAGTCCCGGCCCTCTCAA 2168
 Db 350 GAGGGAGCATGTTGCAAGACCCCGGCATCCAGTGGGGCAGTCCCGGCCCTCTCAA 291
 QY 2169 CTGAGCCGAAAGGCGCCAGCAGGGCAAGACTGAGAGGTGCCCCCAGCCCTGTGCCAGT 2228
 Db 290 CTGAGCCGAAAGGCGCCAGCAGGGCAAGACTGAGAGGTGCCCCCAGCCCTGTGCCAGT 231
 QY 2229 GAAGCCAGCTCCGCCCTCATCTTGAAGCACCAGAGGTGGCCGTGAGGGCCAGATCCCT 2288
 Db 230 GAAGCCAGCTCCGCCCTCATCTTGAAGCACCAGAGGTGGCCGTGAGGGCCAGATCCCT 171
 QY 2289 GCAGCCCTCAGCCCTGGCTCCAGCTGGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2348
 Db 170 GCAGC-CCTCAGCCCTGGCTCCAGCTGGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 112
 QY 2349 AGGACCTCCGGGCAATGCTGTCCCGCTCATCTGAGAGGTGCTCTGGGACCTGCTG 2408
 Db 111 AGGACCTCCGGGCAATGCTGTCCCGCTCATCTGAGAGGTGCTCTGGGACCTGCTG 52
 QY 2409 CCCATTAAAGACTGGTCCAGACCTGTCTGAAAAAAGAAAAAAGAAAAA 2459
 Db 51 CCCATTAAAGACTGGTCCAGACCTGTCTGAAAAAAGAAAAAAGAAAAA 1

RESULT 15
 BX384556
 LOCUS BX384556 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
 DEFINITION cDNA clone CS0DK011YM10 5-PRIME, mRNA sequence.
 ACCESSION BX384556
 VERSION BX384556.1 GI:30440364
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 1201)
 AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of Invitrogen. Contact : Feng Liang Email : liang@lifetech.com URL : <http://fulllength.invitrogen.com/> Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CS0DK011BG05QF1.

FEATURES
 source Location/Qualifiers
 1..1201
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DK011YM10"
 /cell_type="HELA CELLS COT 25-NORMALIZED"
 /clone_lib="HELA"
 /notes="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and Ecor V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
 Query Match 25.3%; Score 624.4; DB 13; Length 1201;
 Best Local Similarity 97.3%; Pred. No. 9.5e-64;
 Matches 651; Conservative 4; Mismatches 11; Indels 3; Gaps 2;
 QY 570 AGCTGTCACTTCCCTCATGATCCCTTGGGCTTGAAGAGCAGGAGCTGGACTGTC 629
 Db 365 AASTTCACTTGTCCCTCATGATCCCTTGGGCTTGAAGAGCAGGAGCTGGACTGTC 424
 QY 630 TACACCTGTAAGGAGCTGATCTCAGTGCATTTGGAGAGACCGCGCTCTCTAGGAGC 689


```
Db      425 TACACCGCTGAGGAGCTGATYTCAGTGCACCTTTGGAGAGGACGGCGCCTCTTACGAGGC 484
QY      690 AGAAATCAGGAGCTGGAGGCCCTGGCGCAGGCGCATGCGGACCCCGAGCGGAATGAGTC 749
Db      485 AGAAATCAGGAGCTGGAGGCCCTGGCGCAGGCGCATGCGGACCCCGAGCGGAATGAGTC 544
QY      750 GGGCCTGGAGCTGCTCACAGCCTATTACAAACAGCTGTGCTTCTCTGGATGGCGCTTCCT 809
Db      545 GGGCCTGGAGCTGCTCACAGCCTATTACAAACAGCTGTGCTTCTCTGGATGGCGCTTCCT 604
QY      810 CACCCCTGCCAGGAGCTCGGGCTCTTCTTCCACTGTAGACTCGCTTACTGGGTCGC 869
Db      605 CACCCCTGCCAGGAGCTCGGGCTCTTCTTCCACTGTAGACTCGCTTACTGGGTCGC 664
QY      870 GGCCAGCAGCGCTGCCCTTGGCCCTTCGAGAAGGSCAGCGTTCTTCAACATCGGTGCCCT 929
Db      665 GGCCAGCAGCGCTGCCCTTGGCCCTTCGAGAAGGSCAGCGTTCTTCAACATCGGTGCCCT 724
QY      930 CCACACGAGATTGGGGCGGCGCAGGACCGCTCCTGACCGAGGGTGGCGCGCGCTAT 989
Db      725 CCACACGAGATTGGGGCGGCGCAGGACCGCTCCTGACCGAGGGTGGCGCGCGCTAT 784
QY      990 GGAGGCTTCCAGAGGCGCGCTGGGGCTTTCAGCCTCCTGAGGGAGAACTTCTCCCATGC 1049
Db      785 GGAGGCTTCCAGAGGCGCGCTGGGGCTTTCAGCCTCCTGAGGGAGAACTTCTCCCATGC 844
QY      1050 GCCGAGCCGAGACATGAGCGCTGGGTCCCTCTGGCACTGGAGCAGCTCATGATGGCCCA 1109
Db      845 GCCGAGCCGAGACATGAGCGCTGGGTCCCTCTGGCACTGGAGCAGCTCATGATGGCCCA 904
QY      1110 GGCCAGGAAATGTGTTTGGGGCTCTCACGACCTGCTCATGGCCCCCAGACTG 1169
Db      905 GGCCAGGAAATGTGTTTGGGGCTCTCACGACCTGCTCATGG-CCCCAAGACTG 963
QY      1170 CCTGGCCAGCTGGCGCTGGCGCAGGCGCGCCAGGTGGCAGCGAGTACAGGCTAGT 1229
Db      964 CTTGGCCAGCTGGCGCTGGCGAGGAGGCG--GCCAGGTGGCAGCGGAGTACAGGCTAGT 1021
QY      1230 GCACGGAC 1238
Db      1022 GCACGGAC 1030
```

Search completed: July 13, 2004, 08:20:59
Job time : 4187.32 secs

his Page Blank (uspto)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.
OM nucleic - nucleic search, using sw model
Run on: July 13, 2004, 06:10:50 ; Search time 1005.06 Seconds
(without alignments)
11973.025 Million cell updates/sec

Title: US-10-697-266-1
Perfect score: 2469
Sequence: 1 tcgcggccagggtggtgcgg.....aaaaaaaaaaaaaaaaaaaaa 2469

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3183909 seqs, 2436941669 residues
Total number of hits satisfying chosen parameters: 6367818

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA.*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2.*
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				
Result No.	Score	Query Match	Length DB ID	Description
1	2469	100.0	2469 13	US-10-697-266-1 Sequence 1, Appli
2	2469	100.0	2469 13	US-10-274-878-1 Sequence 1, Appli
3	2105	85.3	2837 15	US-10-451-207-11 Sequence 11, Appl
4	2082.4	84.3	2738 13	US-10-112-944-74 Sequence 74, Appl
5	957	38.8	1013 15	US-10-023-896-15 Sequence 15, Appl
6	955.4	38.7	1013 9	US-09-925-301-174 Sequence 174, Appl
7	955.4	38.7	1013 15	US-10-023-896-43 Sequence 43, Appl
8	946.8	38.3	1155 9	US-09-833-381-1293 Sequence 1293, Ap
9	674	27.3	1046 13	US-10-112-944-558 Sequence 558, App
10	496.4	20.1	2100 16	US-10-239-607-52 Sequence 52, Appl
11	496.4	20.1	3526 16	US-10-239-607-51 Sequence 51, Appl
12	494.8	20.0	2109 13	US-10-092-900A-351 Sequence 351, App
13	494.8	20.0	3500 16	US-10-108-260A-2397 Sequence 2397, Ap
14	493.2	20.0	2061 9	US-09-895-040A-2 Sequence 2, Appli

15	493.2	20.0	3484 9	US-09-895-040A-1
16	483.8	19.6	3518 10	US-09-814-353-20684
17	463.6	18.8	2310 13	US-10-092-900A-349
18	418.2	16.9	2757 16	US-10-378-029-49
19	368	14.9	19025 13	US-10-697-266-3
20	368	14.9	19025 15	US-10-274-878-3
21	149.6	6.1	2300 16	US-10-108-260A-469
22	108.4	4.4	2925 16	US-10-094-749-256
23	108.4	4.4	4212 13	US-10-425-114-26843
24	108.4	4.4	5200 17	US-10-433-794-21
25	103.6	4.2	599 10	US-03-764-891-244
26	94.6	3.8	439 9	US-09-918-995-4314
27	93.4	3.8	338 13	US-09-815-343-97
28	93.4	3.8	338 13	US-10-097-105-97
29	87	3.5	479 9	US-09-867-701-4120
30	72.4	2.9	500 9	US-09-895-040A-29
31	71	2.9	167 16	US-10-239-607-62
32	70.8	2.9	501 10	US-09-814-353-15780
33	70.8	2.9	553 10	US-09-814-353-17840
34	69.4	2.8	167 9	US-09-895-040A-14
35	69	2.8	799 13	US-10-424-599-133011
36	68.2	2.8	2941 17	US-10-437-963-78091
37	68	2.8	334 13	US-10-027-632-124788
38	68	2.8	334 13	US-10-027-632-124788
39	68	2.8	334 16	US-10-027-632-124788
40	68	2.8	334 16	US-10-027-632-124788
41	68	2.8	2469 13	US-10-697-266-1
42	68	2.8	2469 15	US-10-274-878-1
43	68	2.8	2738 13	US-10-112-944-74
44	66	2.7	2837 13	US-10-451-207-11
45	66	2.7	401 10	US-09-814-353-5166

ALIGNMENTS

RESULT 1
US-10-697-266-1
; Sequence 1, Application US/10697266
; Publication No. US20040067522A1
; GENERAL INFORMATION:
; APPLICANT: RUSCH, Douglas et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001099-CIP-DIV2
; CURRENT APPLICATION NUMBER: US/10/697,266
; PRIOR FILING DATE: 2003-10-31
; PRIOR APPLICATION NUMBER: 10/274,878
; PRIOR FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: 09/849,334
; PRIOR FILING DATE: 2001-05-07
; PRIOR APPLICATION NUMBER: 09/773,371
; PRIOR FILING DATE: 2001-02-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2469
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-697-266-1

Query Match				
Best Local Similarity	100.0%;	Score 2469;	DB 13;	Length 2469;
Matches 2469;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	TCGCGGCCCCAGGTGTTGGTGGCGGCGCCCTAGCCCGGCTCGGAGCGCTCGCGAGCGGCGG	60	
Db	1	TCGCGGCCCCAGGTGTTGGTGGCGGCGCCCTAGCCCGGCTCGGAGCGCTCGCGAGCGGCGG	60	
Qy	61	GCTGCTGACCCCCAGGAGGACCCCCAGCGGCGGCGGCGATGATCTCTGAGGAGGAGG	120	
Db	61	GCTGCTGACCCCCAGGAGGACCCCCAGCGGCGGCGGCGATGATCTCTGAGGAGGAGG	120	

Db	1501	ATCCTGGGACGAGGAGGGCTGCGGCTGCACGCCCTGTGCGCGCTCTCTGGCGAGGTG	1560	APPLICANT: BURFORD, Neil
Qy	1561	GACCTCTTGGGCTGTGATCTCCAGACGCTGACGGTCTCACTGGCCAAGTATGGGAG	1620	APPLICANT: DING, Li
Db	1561	GACCTCTTGGGCTGTGATCTCCAGACGCTGACGGTCTCACTGGCCAAGTATGGGAG	1620	APPLICANT: YUE, Henry
Qy	1621	CTGACCGTGAGATGACTTCTGTGAGGCTGCGAGGCCCGGACATCCAGCCCTAGAGCC	1680	APPLICANT: THORNTON, Michael B.
Db	1621	CTGACCGTGAGATGACTTCTGTGAGGCTGCGAGGCCCGGACATCCAGCCCTAGAGCC	1680	APPLICANT: CHANLA, Narinder K.
Qy	1681	CACAGAAGCCAGAGCCAGATGCCACGCTGTCCAGGGGAAGGGGCTGACATCTTC	1740	APPLICANT: GANDHI, Ameena R.
Db	1681	CACAGAAGCCAGAGCCAGATGCCACGCTGTCCAGGGGAAGGGGCTGACATCTTC	1740	APPLICANT: ARVIZU, Chandra S.
Qy	1741	CATCGGTGGGCCCCCTGTCTGTGTTCTAGCCAGAACCCGGTGGCGCTGTGGGCC	1800	APPLICANT: BAUGHN, Mariah R.
Db	1741	CATCGGTGGGCCCCCTGTCTGTGTTCTAGCCAGAACCCGGTGGCGCTGTGGGCC	1800	APPLICANT: SWARNAKAR, Anita
Qy	1801	GTCACCTGACCCGAGAGAGGGCGGCTTTGGCTCAGCGTTCGGGGAGACTCGCCTGTC	1860	APPLICANT: DUGGAN, Brendan M.
Db	1801	GTCACCTGACCCGAGAGAGGGCGGCTTTGGCTCAGCGTTCGGGGAGACTCGCCTGTC	1860	APPLICANT: LU, Dying Aina M.
Qy	1861	CTCATCGTCCGCTCATTTCCAGGAGCCAGCCGCGCGCTGCTTGAAGGAGGGCGAC	1920	APPLICANT: THANGAVELU, Kavitha
Db	1861	CTCATCGTCCGCTCATTTCCAGGAGCCAGCCGCGCGCTGCTTGAAGGAGGGCGAC	1920	APPLICANT: WARREN, Bridget A.
Qy	1921	TACATTGTGTCAGTAAATGGGACGCTATGAGGTGGTGAGACACGCGGAGGTGTGACG	1980	APPLICANT: TANG, Y. Tom
Db	1921	TACATTGTGTCAGTAAATGGGACGCTATGAGGTGGTGAGACACGCGGAGGTGTGACG	1980	APPLICANT: KHAN, Farrah A.
Qy	1981	GAGCTGAAGCTGCGGAGAGAGGGCGCGCAGCTGCGAGTGGTGTGCTGCTGCCCGACG	2040	APPLICANT: YAO, Monique G.
Db	1981	GAGCTGAAGCTGCGGAGAGAGGGCGCGCAGCTGCGAGTGGTGTGCTGCTGCCCGACG	2040	APPLICANT: EMERLING, Brooke M.
Qy	2041	TCTAGACTGCCAGCTTTGGGAGCCGCGCGGCTGCTGCTGGGCCCGCCAGGCGGCTTCTA	2100	APPLICANT: TITLE OF INVENTION: INTRACELLULAR SIGNALING MOLECULES
Db	2041	TCTAGACTGCCAGCTTTGGGAGCCGCGCGGCTGCTGCTGGGCCCGCCAGGCGGCTTCTA	2100	FILE REFERENCE: PF-0868 USN
Qy	2101	AGGAGCCAGAGGAGGATGTTGTCAGACCCCGGCTCCAGCTGGGCGGCTCCCGGCC	2160	CURRENT APPLICATION NUMBER: US/10/451,207
Db	2101	AGGAGCCAGAGGAGGATGTTGTCAGACCCCGGCTCCAGCTGGGCGGCTCCCGGCC	2160	CURRENT FILING DATE: 2003-06-18
Qy	2161	CTCCTCAACTGGAGCCGAAAGCCAGCAGGCGCAAGACTGAGGCTGCCCCCAGCCCTGT	2220	PRIOR APPLICATION NUMBER: PCT/US01/50315
Db	2161	CTCCTCAACTGGAGCCGAAAGCCAGCAGGCGCAAGACTGAGGCTGCCCCCAGCCCTGT	2220	PRIOR FILING DATE: 2001-12-19
Qy	2221	GCCCCAGTGAAGCAGCTCCGCGCTCATCTTGAAGCACCCAGGGTGGCGTGAAGGCCA	2280	PRIOR APPLICATION NUMBER: US 60/257,804
Db	2221	GCCCCAGTGAAGCAGCTCCGCGCTCATCTTGAAGCACCCAGGGTGGCGTGAAGGCCA	2280	PRIOR FILING DATE: 2000-12-21
Qy	2281	GGATCCCTGCACGCCCTCAGCCCTGCTCCAGCTGGCAGCAGCACCGAGCATGCCCTCC	2340	PRIOR APPLICATION NUMBER: US 60/260,102
Db	2281	GGATCCCTGCACGCCCTCAGCCCTGCTCCAGCTGGCAGCAGCACCGAGCATGCCCTCC	2340	NUMBER OF SEQ ID NOS: 18
Qy	2341	CCACCCAGAGGACCTCCGGGCAATGCTGCTCCGCTCATCTCTGAGAGCTCCCTCGGCA	2400	SOFTWARE: PERL Program
Db	2341	CCACCCAGAGGACCTCCGGGCAATGCTGCTCCGCTCATCTCTGAGAGCTCCCTCGGCA	2400	SEQ ID NO 11
Qy	2401	CCTGCTGCCCATTAAGACTGTGTGAGCTGTCTGAGGAGGAGGAGGAGGAGGAGGAGG	2460	LENGTH: 2837
Db	2401	CCTGCTGCCCATTAAGACTGTGTGAGCTGTCTGAGGAGGAGGAGGAGGAGGAGGAGG	2460	TYPE: DNA
Qy	2461	AAAAAAAAA 2469		ORGANISM: Homo sapiens
Db	2461	AAAAAAAAA 2469		NAME/KEY: misc feature
RESULT 3				OTHER INFORMATION: Incyte ID No. US20040038267A1 5560161CB1
US-10-451-207-11				US-10-451-207-11
; Sequence 11, Application US/10451207				Query Match 85.3%; Score 2105; DB 13; Length 2837;
; Publication No. US20040038267A1				Best Local Similarity 100.0%; Pred. No. 0;
; GENERAL INFORMATION:				Matches 2116; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
; APPLICANT: INCYTE GENOMICS, INC.				

QY 801 GCGCTTCTCACCCTGCCAGAGCCTCGGGCTCTTCTTCCACTGTACGACTCGCTTAC 860
Db |||||
661 GCGCTTCTCACCCTGCCAGAGCCTCGGGCTCTTCTTCCACTGTACGACTCGCTTAC 720
QY 861 TGGGGTCCCGGGCCAGCAGCTGCGCTGCGCTTCCAGAAAGGCGAGGGTCTCTTCAACAT 920
Db |||||
721 TGGGGTCCCGGGCCAGCAGCTGCGCTGCGCTTCCAGAAAGGCGAGGGTCTCTTCAACAT 780
QY 921 CGGTGCTTCCACAGCAGATTGGGGCGCGCCAGAGACCGCTCTCTGACCGAGGGTGCCTCG 980
Db |||||
781 CGGTGCTTCCACAGCAGATTGGGGCGCGCCAGAGACCGCTCTCTGACCGAGGGTGCCTCG 840
QY 981 CGGCTATAGAGGCTTCCAGAGGGCGCTGGGGCTTTCAGCCTCTGAGGGAGAACTT 1040
Db |||||
841 CGGCTATAGAGGCTTCCAGAGGGCGCTGGGGCTTTCAGCCTCTGAGGGAGAACTT 900
QY 1041 CTCCCATGCGCGAGCCACAGACATGAGCGCTGCGTCCCTCTGCGCACTGGAGCAGCTCAT 1100
Db |||||
901 CTCCCATGCGCGAGCCACAGACATGAGCGCTGCGTCCCTCTGCGCACTGGAGCAGCTCAT 960
QY 1101 GATGCCCCAGGCCACAGGAATGTGTTGAGGGCTCTCAACACCTGCTCCATGGCCCC 1160
Db |||||
961 GATGCCCCAGGCCACAGGAATGTGTTGAGGGCTCTCAACACCTGCTCCATGGCCCC 1020
QY 1161 CCAAGACTGCTGCGCCAGCTGCGCTGCGCGCAGAGGCGCGCCAGGTGGCAGCGAGTA 1220
Db |||||
1021 CCAAGACTGCTGCGCCAGCTGCGCTGCGCGCAGAGGCGCGCCAGGTGGCAGCGAGTA 1080
QY 1221 CAGGCTAGTGACCGGACCATGCGCCAGCCACCGCTCCACGACTACGTGCTGTCTCGTG 1280
Db |||||
1081 CAGGCTAGTGACCGGACCATGCGCCAGCCACCGCTCCACGACTACGTGCTGTCTCGTG 1140
QY 1281 GACTGCTGTGTGATGCAAGGCGAGTACTTCGCTCCCTGGCCCACTACCACTGATG 1340
Db |||||
1141 GACTGCTGTGTGATGCAAGGCGAGTACTTCGCTCCCTGGCCCACTACCACTGATG 1200
QY 1341 CATGSCCTCTGCGAGCGCTCCCGAGCAGCAGGAGAGCTCCCGCAGCAGAGAGGT 1400
Db |||||
1201 CATGSCCTCTGCGAGCGCTCCCGAGCAGCAGGAGAGCTCCCGCAGCAGAGAGGT 1260
QY 1401 CTTCCTGAGCCCTCCACTCTCTTAAGCCCGCGAGGCGCTGTGTCGCGCAGGAGCTGGA 1460
Db |||||
1261 CTTCCTGAGCCCTCCACTCTCTTAAGCCCGCGAGGCGCTGTGTCGCGCAGGAGCTGGA 1320
QY 1461 GGAGCGCAGGAGCTTGGCAAGGCAACCTGAAGCGTGCCATCTCTGGGCGCAGGAGGAG 1520
Db |||||
1321 GGAGCGCAGGAGCTTGGCAAGGCAACCTGAAGCGTGCCATCTCTGGGCGCAGGAGGAG 1380
QY 1521 GCTCGGCTGACCGCTGTGCGCGCTCTGCGGAGTGACCTGCTTCGGGCTGTGAT 1580
Db |||||
1381 GCTCGGCTGACCGCTGTGCGCGCTCTGCGGAGTGACCTGCTTCGGGCTGTGAT 1440
QY 1581 CTCCAGACGCTGACGCGCTCACTGGGCAAGTATGCGGAGCTGCACCGTGAGGATGACTT 1640
Db |||||
1441 CTCCAGACGCTGACGCGCTCACTGGGCAAGTATGCGGAGCTGCACCGTGAGGATGACTT 1500
QY 1641 CTGTGAGGCTGCGAGGCGCCCGGACATCCAGCTTAAGACCCACAGAGCCAGAGGCGCAG 1700
Db |||||
1501 CTGTGAGGCTGCGAGGCGCCCGGACATCCAGCTTAAGACCCACAGAGCCAGAGGCGCAG 1560
QY 1701 GATGCCAGCTGTCCAGGGAGGGGCGCTGACATCTTCATCGGCTGGGGCGCCCTGTC 1760
Db |||||
1561 GATGCCAGCTGTCCAGGGAGGGGCGCTGACATCTTCATCGGCTGGGGCGCCCTGTC 1620
QY 1761 TGTGTTCTCAGCCAAAGAACCGGTGCGGCTGGTGGGGCGCGCTGCACCTGACCCGAGGAGA 1820
Db |||||
1621 TGTGTTCTCAGCCAAAGAACCGGTGCGGCTGGTGGGGCGCGCTGCACCTGACCCGAGGAGA 1680
QY 1821 GGGCGGCTTGGGCTCAGCTTCTGGGAGACTCGCCTGTCTCATCGCTGCGCTCATTC 1880
Db |||||
1681 GGGCGGCTTGGGCTCAGCTTCTGGGAGACTCGCCTGTCTCATCGCTGCGCTCATTC 1740

QY 1881 AGGGAGCCAGGCGCGCGGCTTGGCTGAAGGAGGCGGACTACTATTGTCTCAGTGAATGG 1940
Db |||||
1741 AGGGAGCCAGGCGCGCGGCTTGGCTGAAGGAGGCGGACTACTATTGTCTCAGTGAATGG 1800
QY 1941 GCAGCCATGCAAGTGGTGGAGACACGCGAGGTGGTACGAGCTGAAGGCTGCGGAGGA 2000
Db |||||
1801 GCAGCCATGCAAGTGGTGGAGACACGCGAGGTGGTACGAGCTGAAGGCTGCGGAGGA 1860
QY 2001 GCGCGGCGCCAGCTGCGAGTGGTGTGCTGCTGCCAGCTCTAGACTGCCAGCTTGG 2060
Db |||||
1861 GCGCGGCGCCAGCTGCGAGTGGTGTGCTGCTGCCAGCTCTAGACTGCCAGCTTGG 1920
QY 2061 GCGCGGCGCCAGCTGCGCTCTGCTGGGCGCCAGGCGGCTTCTAAGGAGCCAGAGGAGCATGG 2120
Db |||||
1921 GCGCGGCGCCAGCTGCGCTCTGCTGGGCGCCAGGCGGCTTCTAAGGAGCCAGAGGAGCATGG 1980
QY 2121 TTGCAAGACCCCGGACTCACTGGGCGCAGTCCCCCGGCGCTCTCTCAACTGAGAGCGGAAA 2180
Db |||||
1981 TTGCAAGACCCCGGACTCACTGGGCGCAGTCCCCCGGCGCTCTCTCAACTGAGAGCGGAAA 2040
QY 2181 GCGCCAGCAGGCGCAAGACTGGAGGCTGCGCCCGCAGCGCTGCCCCAGGCTTGTGCCCGAGTGAAGCCAGCTCC 2240
Db |||||
2041 GCGCCAGCAGGCGCAAGACTGGAGGCTGCGCCCGCAGCGCTTGTGCCCGAGTGAAGCCAGCTCC 2100
QY 2241 GCGCTCATCTTGAAGCACCCAGAGTGGCGCTGAGGGCGCAGGATCCCTGCGAGCGCTCAG 2300
Db |||||
2101 GCGCTCATCTTGAAGCACCCAGAGTGGCGCTGAGGGCGCAGGATCCCTGCGAGCG -CCTCAG 2159
QY 2301 CCCTGGCTCAGCTGGGCAAGCACCGAGCATGCCCTCCCGCAGGAGCTCCCGG 2360
Db |||||
2160 CCCTGGCTCAGCTGGGCAAGCACCGAGCATGCCCTCCCGCAGGAGCTCCCGG 2219
QY 2361 CAATGCTGTGCTCCGCTCATGCTGGAGGCTGCTGCGGCACTGCTGCGGCACTTAAAGAC 2420
Db |||||
2220 CAATGCTGTGCTCCGCTCATGCTGGAGGCTGCTGCGGCACTGCTGCGGCACTTAAAGAC 2279
QY 2421 TGGTCAGACTGTCTGA 2437
Db |||||
2280 TGGTCAGACTGTCTGA 2296

RESULT 4
US-10-112-944-74
; Sequence 74, Application US/10112944
; Publication No. US20040048249A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Yang, Yonghong
; APPLICANT: Weng, Gezhi
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Xue, Aidong J.
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Wehrman, Tom
; APPLICANT: Ghosh, Malabika
; APPLICANT: Wang, Dunrui
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wang, Zhiwei
; TITLE OF INVENTION: No. US20040048249A1el Nucleic Acids and
; TITLE OF INVENTION: Secreted Polypeptides
; FILE REFERENCE: 805A
; CURRENT APPLICATION NUMBER: US/10/112.944
; CURRENT FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: US 09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 09/491,404
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: US 09/496,914
; PRIOR FILING DATE: 2000-02-03
; PRIOR APPLICATION NUMBER: US 09/515,126
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: US 09/519,705
; PRIOR FILING DATE: 2000-03-07

; PRIOR APPLICATION NUMBER: US 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: US 09/552,929
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: US 09/577,408
; PRIOR FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 924
; SOFTWARE: pt_FL_genes Version 5.0
; SEQ ID NO 74
; LENGTH: 2738
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (103)..(2109)
US-10-112-944-74

Query Match 84.3%; Score 2082.4; DB 13; Length 2738;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 2116; Conservative 0; Mismatches 16; Indels 4; Gaps 2;
QY 321 GGGCTGTGACTCCCTGACGAGATCCAGTGGCGCCAGCTGCGAGCGCGAGGCGCCAGAT 380
Db |||||
QY 162 GGGCTGTGACTCCCTGACGAGATCCAGTGGCGCCAGCTGCGAGCGCGAGGCGCCAGAT 221
Db |||||
QY 381 TCACGAGCAGATTGACAAAGGAGCTGCAGATGCGGCGGCGCTGAGAACCTCTACAGAGC 440
Db |||||
QY 222 TCACGAGCAGATTGACAAAGGAGCTGCAGATGCGGCGGCGCTGAGAACCTCTACAGAGC 281
Db |||||
QY 441 CACGAGCAACACCGGCTGAGAGAGACGGTTCGCCCTGGAGCTGAGCTACGTCACATCCAA 500
Db |||||
QY 282 CACGAGCAACACCGGCTGAGAGAGAGCGTTCGCCCTGGAGCTGAGCTACGTCACATCCAA 341
Db |||||
QY 501 CTTGCGAGCTCTGAAGAGAGCTGGAGAGCTCAGCGGTGGCGTGGAGCCCTGGCGCGCA 560
Db |||||
QY 342 CTTGCGAGCTCTGAAGAGAGCTGGAGAGCTCAGCGGTGGCGTGGAGCCCTGGCGCGCA 401
Db |||||
QY 561 TGGGAGCGAGCTGTCACTGTCCCATGATCCCGCTGGGCTGAAGGAGACCAAGGAGCT 620
Db |||||
QY 402 TGGGAGCGAGCTGTCACTGTCCCATGATCCCGCTGGGCTGAAGGAGACCAAGGAGCT 461
Db |||||
QY 621 GGAAGTGTCTACACCGCTGAAGAGCTGTATCTCAGTGCATCTTTGGAGAGACCGCGCTC 680
Db |||||
QY 462 GGAAGTGTCTACACCGCTGAAGAGCTGTATCTCAGTGCATCTTTGGAGAGACCGCGCTC 521
Db |||||
QY 681 CTACGAGCGAGAAATCAGGAGCTGAGGAGCTTGGCGGAGCCATGCGGAGCCCGCCAGCG 740
Db |||||
QY 522 CTACGAGCGAGAAATCAGGAGCTGAGGAGCTTGGCGGAGCCATGCGGAGCCCGCCAGCG 581
Db |||||
QY 741 GAATGAGTTCGGGCTGGAGCTGTCAAGGCTATTACAAACAGCTGTGCTTCCTGGATGC 800
Db |||||
QY 582 GAATGAGTTCGGGCTGGAGCTGTCAAGGCTATTACAAACAGCTGTGCTTCCTGGATGC 641
Db |||||
QY 801 GCGCTTCTTCCCTCAGGAGGCTCGGGCTCTTCTTCCACTGGTACGACTCGCTTAC 860
Db |||||
QY 642 GCGCTTCTTCCCTCAGGAGGCTCGGGCTCTTCTTCCACTGGTACGACTCGCTTAC 701
Db |||||
QY 861 TGGGGTCCCGGCGAGGAGCTGCGCTTGGGCTTCGAGAGGCGAGCTTCTTTCACAT 920
Db |||||
QY 702 TGGGGTCCCGGCGAGGAGCTGCGCTTGGGCTTCGAGAGGCGAGCTTCTTTCACAT 761
Db |||||
QY 921 CGGTGCGCTCCACAGCAGATTGGGCGCGCAGGAGCGCTTCCTGACCGAGGCTGCGCG 980
Db |||||
QY 762 CGGTGCGCTCCACAGCAGATTGGGCGCGCAGGAGCGCTTCCTGACCGAGGCTGCGCG 821
Db |||||
QY 981 CCGCGCTATGAGGCTTCCAGAGGCGCGTGGGCGCTTACGCTCTCTGAGGGAGAACTT 1040
Db |||||
QY 822 CCGCGCTATGAGGCTTCCAGAGGCGCGTGGGCGCTTACGCTCTCTGAGGGAGAACTT 881
Db |||||
QY 1041 CTCCTATGCGGAGCGCCAGACATGAGGCTGCGCTTCTGCGACTGGAGCAGCTCAT 1100
Db |||||
QY 882 CTCCTATGCGGAGCGCCAGACATGAGGCTGCGCTTCTGCGACTGGAGCAGCTCAT 941
Db |||||

QY 1101 GATGGCCAGGCGCCAGGAATGTGTGTTGAGGCGCTCTTCCACCTGCTTCCATGGCCCC 1160
Db |||||
QY 942 GATGGCCAGGCGCCAGGAATGTGTGTTGAGGCGCTCTTCCACCTGCTTCCATGGCCCC 1001
Db |||||
QY 1161 CCAAGACTGCTGCGCCAGCTGCGCTGCGGAGGCGCGCCAGGCTGGAGCCAGGTA 1220
Db |||||
QY 1002 CCAAGACTGCTGCGCCAGCTGCGCTGCGGAGGCGCGCCAGGCTGGAGCCAGGTA 1061
Db |||||
QY 1221 CAGGCTAGTGCAACGAGACCATGSCCAGCCACCCGTCACAGCTAGTGCTCTCTCTG 1280
Db |||||
QY 1062 CAGGCTAGTGCAACGAGACCATGSCCAGCCACCCGTCACAGCTAGTGCTCTCTCTG 1121
Db |||||
QY 1281 GACTGCGCTGTGTCATGTCAAGGCGGAGTACTTCCGCTCTCTGCGCCACTACCACTAGC 1340
Db |||||
QY 1122 GACTGCGCTGTGTCATGTCAAGGCGGAGTACTTCCGCTCTCTGCGCCACTACCACTAGC 1181
Db |||||
QY 1341 CATGGCCCTGTGCAACGAGCTGCGGAGGCGGAGGAGCTGCGCCAGCAGGAGCT 1400
Db |||||
QY 1182 CATGGCCCTGTGCAACGAGCTGCGGAGGAGGAGCTGCGCCAGCAGGAGCT 1238
Db |||||
QY 1401 CTTCTGCGAGCGCCCGACCTCTTAAAGCCCGAGGCGCTGTGCTGCGCGAGGAGCTGGA 1460
Db |||||
QY 1239 CTTCTGCGAGCGCCCGACCTCTTAAAGCCCGAGGCGCTGTGCTGCGCGAGGAGCTGGA 1298
Db |||||
QY 1461 GGAGCGCAGGAGCTTGGCAAGGACACCTGAAAGCGTGCCATCTTGGGGGAGGAGGAGC 1520
Db |||||
QY 1299 GGAGCGCAGGAGCTTGGCAAGGACACCTGAAAGCGTGCCATCTTGGGGGAGGAGGAGC 1358
Db |||||
QY 1521 GCTGCGGCTGCAACGCGCTGTGCGGCTGCTGCGGAGGAGGAGCTGCTTGGGCTGTGAT 1580
Db |||||
QY 1359 GCTGCGGCTGCAACGCGCTGTGCGGCTGCTGCGGAGGAGGAGCTGCTTGGGCTGTGAT 1418
Db |||||
QY 1581 CTCGCCAGAGCTGCAACGCGCTGCTGCGGAGGAGGAGCTGCGGAGGAGGAGCTGAT 1640
Db |||||
QY 1419 CTCGCCAGAGCTGCAACGCGCTGCTGCGGAGGAGGAGCTGCGGAGGAGGAGCTGAT 1478
Db |||||
QY 1641 CTTGAGGCTGCGGAGCGCGGAGCTGCAACCTGAAAGCGTGCCATCTTGGGGGAGGAGGAGC 1700
Db |||||
QY 1479 CTTGAGGCTGCGGAGCGCGGAGCTGCAACCTGAAAGCGTGCCATCTTGGGGGAGGAGGAGC 1538
Db |||||
QY 1701 GATGCCAGCGCTGCTCCAGGAGGAGGAGCTGCACTTCCATGCGGCTGCGGCGCTGTC 1760
Db |||||
QY 1539 GATGCCAGCGCTGCTCCAGGAGGAGGAGCTGCACTTCCATGCGGCTGCGGCGCTGTC 1598
Db |||||
QY 1761 TGTGTTCTCAGCAAGAAACCGGTGGGCGCTGCTGCGGCGCTGCACTTCCATGCGGCTGCGGCGCTGTC 1820
Db |||||
QY 1599 TGTGTTCTCAGCAAGAAACCGGTGGGCGCTGCTGCGGCGCTGCACTTCCATGCGGCTGCGGCGCTGTC 1658
Db |||||
QY 1821 GGGCGGCTTGTGCTTCCAGCTTCCAGGAGGAGCTGCGGCGCTGCACTTCCATGCGGCTGCGGCGCTGTC 1880
Db |||||
QY 1659 GGGCGGCTTGTGCTTCCAGCTTCCAGGAGGAGCTGCGGCGCTGCACTTCCATGCGGCTGCGGCGCTGTC 1718
Db |||||
QY 1881 AGGAGCGAGCGCGCGGCTGCGGCTGAGGAGGAGGAGCTGCACTTCCATGCGGCTGCGGCGCTGTC 1940
Db |||||
QY 1719 AGGAGCGAGCGCGCGGCTGCGGCTGAGGAGGAGGAGCTGCACTTCCATGCGGCTGCGGCGCTGTC 1778
Db |||||
QY 1941 GCAGCCATGAGGCTGCTGAGAGACACCGGAGGAGGAGCTGCACTTCCATGCGGCTGCGGCGCTGTC 2000
Db |||||
QY 1779 GCAGCCATGAGGCTGCTGAGAGACACCGGAGGAGGAGCTGCACTTCCATGCGGCTGCGGCGCTGTC 1838
Db |||||
QY 2001 GGGCGGCGCGGAGCTGAGGCTGCTGCGGCGCTGAGAGCTGAGAGCTGCGGCGCTGTC 2060
Db |||||
QY 1839 GGGCGGCGCGGAGCTGAGGCTGCTGCGGCGCTGAGAGCTGAGAGCTGCGGCGCTGTC 1898
Db |||||
QY 2061 GGGCGGCGCGGAGCTGAGGCTGCTGCGGCGCTGAGAGCTGAGAGCTGCGGCGCTGTC 2120
Db |||||
QY 1899 GGGCGGCGCGGAGCTGAGGCTGCTGCGGCGCTGAGAGCTGAGAGCTGCGGCGCTGTC 1958
Db |||||
QY 2121 TTGCAAGACCGCGGAGCTGAGGCTGCTGCGGCGCTTCCATGCGGCTGCGGCGCTGTC 2180
Db |||||
QY 1959 TTGCAAGACCGCGGAGCTGAGGCTGCTGCGGCGCTTCCATGCGGCTGCGGCGCTGTC 2018
Db |||||
QY 2181 GGGCGGCGCGGAGCTGAGGCTGCTGCGGCGCTTCCATGCGGCTGCGGCGCTGTC 2240
Db |||||

Db 27 GGCAGGCAACCTGAGCGTGCATCTCTGGGAGGAGGAGCGCTGGCGTGCAGCC 86
QY 1537 CTGTGCGCGTCTCTGCGAGGTGGACCTGCTTTCGGGCTGTGATCTCCAGACGCTGCAG 1596
Db 87 CTGTGCGCGTCTCTGCGAGGTGGACCTGCTTTCGGGCTGTGATCTCCAGACGCTGCAG 146
QY 1597 CGCTCACTGGCCTAAGTATGCGGAGCTGCACCTGAGGATGACTTCTGTGAGGCTGCCGAG 1656
Db 147 CGCTCACTGGCCTAAGTATGCGGAGCTGCACCTGAGGATGACTTCTGTGAGGCTGCCGAG 206
QY 1657 GCCCGGACATCAGCCTTAAGACCCACCAAGCCAGAGCCAGGATGCCACGCTGTCC 1716
Db 207 GCCCGGACATCAGCCTTAAGACCCACCAAGCCAGAGCCAGGATGCCACGCTGTCC 266
QY 1717 CAGGGAGAGGGGCTGACATCTTCCATTCGCTGCGGCGCCCTGTCTGTGTTCTCAGCCAG 1776
Db 267 CAGGGAGAGGGGCTGACATCTTCCATTCGCTGCGGCGCCCTGTCTGTGTTCTCAGCCAG 326
QY 1777 AACGGTGGCGGCTGGTGGGCGCGCTCCACTGACCCGAGGAGGCGGCTTTGGCTC 1836
Db 327 AACGGTGGCGGCTGGTGGGCGCGCTCCACTGACCCGAGGAGGCGGCTTTGGCTC 386
QY 1837 ACCTTTCGGGAGACTCGCTGTCTCATCGCTGCCCTCATTCAGGAGCCAGGCGCG 1896
Db 387 ACCTTTCGGGAGACTCGCTGTCTCATCGCTGCCCTCATTCAGGAGCCAGGCGCG 446
QY 1897 GCGGCTGGCTGAAGAGGCGGCTAATTTGTGTGAGTGAATGGGAGCCAGGCTG 1956
Db 447 GCGGCTGGCTGAAGAGGCGGCTAATTTGTGTGAGTGAATGGGAGCCAGGCTG 506
QY 1957 TGGAGACACGCGAGGTGGTACGGAGCTGAAGCTCGGAGAGGCGGCGCGAGCTG 2016
Db 507 TGGAGACACGCGAGGTGGTACGGAGCTGAAGCTCGGAGAGGCGGCGCGAGCTG 566
QY 2017 CAGGTGTGTGCTGTGCTGCCAGCTTAGAGCTGCTTGGGAGCCCGCGCGCTC 2076
Db 567 CAGGTGTGTGCTGTGCTGCCAGCTTAGAGCTGCTTGGGAGCCCGCGCGCTC 626
QY 2077 CTGCTGGGCGCCAGGGGCTTTAAGAGCCAGGAGGAGCATGGTTGCAAGACCCGCGCA 2136
Db 627 CTGCTGGGCGCCAGGGGCTTTAAGAGCCAGGAGGAGCATGGTTGCAAGACCCGCGCA 686
QY 2137 TCCACGTGGCGCTGCTCCCGCGCTCTCAACTGGAGCCGAAAGGCCAGCGGCAAG 2196
Db 687 TCCACGTGGCGCTGCTCCCGCGCTCTCAACTGGAGCCGAAAGGCCAGCGGCAAG 746
QY 2197 ACTGGAGGCTGCCCCCAGCCCTGTGCCCCAGTGAAGCCAGCTCCGCTCATCTTGAAG 2256
Db 747 ACTGGAGGCTG-CCCCAGCCCTGTGCCCCAGTGAAGCCAGCTCCGCTCATCTTGAAG 805
QY 2257 CACCAGGTGGCGGTGAGGCGCAGGATCCCTGACGCGCTCAGCCCTGGCTCCAGCTGG 2316
Db 806 CACCAGGTGGCGGTGAGGCGCAGGATCCCTGACG-CTTCAAGCCCTGGCTCCAGCTGG 864
QY 2317 CAGCAAGCAGCAGCATGCTCCCTCCCAAGAGGAGCTCCGCGCAATGCTGTCCGCGC 2376
Db 865 CAGCAAGCAGCAGCATGCTCCCTCCCAAGAGGAGCTCCGCGCAATGCTGTCCGCGC 924
QY 2377 TCATGCTGGAGGCTGCTTGGGCGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2436
Db 925 TCATGCTGGAGGCTGCTTGGGCGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 984
QY 2437 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2465
Db 985 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1013

RESULT 7
US-10-023-896-43
; Sequence 43, Application US/10023896
; Publication No. US2003002776A1
; GENERAL INFORMATION:

; APPLICANT: Victor Roschke
; TITLE OF INVENTION: 29 Human Cancer Associated Proteins
; FILE REFERENCE: PA004P1
; CURRENT APPLICATION NUMBER: US/10/023,896
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: unassigned
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: PCT/US00/23794
; PRIOR FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/152,296
; PRIOR FILING DATE: 1999-09-03
; PRIOR APPLICATION NUMBER: 60/158,003
; PRIOR FILING DATE: 1999-10-06
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 43
; LENGTH: 1013
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-023-896-43
Query Match 38.7%; Score 955.4; DB 15; Length 1013;
Best Local Similarity 99.2%; Pred. No. 5.7e-222;
Matches 981; Conservative 0; Mismatches 6; Indels 2; Gaps 2;
QY 1477 GGCAGGCAACCTGAGCGTGCATCTCTGGGAGGAGGAGCGCTGGCGTGCAGCC 1536
Db 27 GGCAGGCAACCTGAGCGTGCATCTCTGGGAGGAGGAGCGCTGGCGTGCAGCC 86
QY 1537 CTGTGCGCGTCTCTGCGAGGTGGACCTGCTTTCGGGCTGTGATCTCCAGACGCTGCAG 1596
Db 87 CTGTGCGCGTCTCTGCGAGGTGGACCTGCTTTCGGGCTGTGATCTCCAGACGCTGCAG 146
QY 1597 CGCTCACTGGCCTAAGTATGCGGAGCTGCACCTGAGGATGACTTCTGTGAGGCTGCCGAG 1656
Db 147 CGCTCACTGGCCTAAGTATGCGGAGCTGCACCTGAGGATGACTTCTGTGAGGCTGCCGAG 206
QY 1657 GCCCGGACATCAGCCTTAAGACCCACCAAGCCAGAGCCAGGATGCCACGCTGTCC 1716
Db 207 GCCCGGACATCAGCCTTAAGACCCACCAAGCCAGAGCCAGGATGCCACGCTGTCC 266
QY 1717 CAGGGAGAGGGGCTGACATCTTCCATTCGCTGCGGCGCCCTGTCTGTGTTCTCAGCCAG 1776
Db 267 CAGGGAGAGGGGCTGACATCTTCCATTCGCTGCGGCGCCCTGTCTGTGTTCTCAGCCAG 326
QY 1777 AACGGTGGCGGCTGGTGGGCGCGCTCCACTGACCCGAGGAGGCGGCTTTGGCTC 1836
Db 327 AACGGTGGCGGCTGGTGGGCGCGCTCCACTGACCCGAGGAGGCGGCTTTGGCTC 386
QY 1837 ACCTTTCGGGAGACTCGCTGTCTCATCGCTGCCCTCATTCAGGAGCCAGGCGCG 1896
Db 387 ACCTTTCGGGAGACTCGCTGTCTCATCGCTGCCCTCATTCAGGAGCCAGGCGCG 446
QY 1897 GCGGCTGGCTGAAGAGGCGGCTAATTTGTGTGAGTGAATGGGAGCCAGGCTG 1956
Db 447 GCGGCTGGCTGAAGAGGCGGCTAATTTGTGTGAGTGAATGGGAGCCAGGCTG 506
QY 1957 TGGAGACACGCGAGGTGGTACGGAGCTGAAGCTCGGAGAGGCGGCGCGAGCTG 2016
Db 507 TGGAGACACGCGAGGTGGTACGGAGCTGAAGCTCGGAGAGGCGGCGCGAGCTG 566
QY 2017 CAGGTGTGTGCTGTGCTGCCAGCTTAGAGCTGCTTGGGAGCCCGCGCGCTC 2076
Db 567 CAGGTGTGTGCTGTGCTGCCAGCTTAGAGCTGCTTGGGAGCCCGCGCGCTC 626
QY 2077 CTGCTGGGCGCCAGGGGCTTTAAGAGCCAGGAGGAGCATGGTTGCAAGACCCGCGCA 2136
Db 627 CTGCTGGGCGCCAGGGGCTTTAAGAGCCAGGAGGAGCATGGTTGCAAGACCCGCGCA 686
QY 2137 TCCACGTGGCGCTGCTCCCGCGCTCTCAACTGGAGCCGAAAGGCCAGCGGCAAG 2196
Db 687 TCCACGTGGCGCTGCTCCCGCGCTCTCAACTGGAGCCGAAAGGCCAGCGGCAAG 746

QY 2197 ACTGAGGCTGCCCGCAGCCCTGTGCCCCAGTGAAGCCAGCTCCGCCCTCATCTTTGAAG 2256
DB |||||
747 ACTGAGGCTG-CCCCAGGCTGTGCCCCAGTGAAGCCAGCTCCGCCCTCATCTTTGAAG 805
QY 2257 CACCCAGGCTGGCGCTGAGGGCCAGGATCCCTGACGCCCTCAGCCCTCGCTCCAGCTGG 2316
DB |||||
806 CACCCAGGCTGGCGCTGAGGGCCAGGATCCCTGACG-CCTCAGCCCTCGCTCCAGCTGG 864
QY 2317 CAGCAAGCACCCAGGATGCTCCCTCCACCCAGAGGACCTCCGGGCAATGCCCTGTCCCGCC 2376
DB |||||
865 CAGCAAGCACCCAGGATGCTCCCTCCACCCAGAGGACCTCCGGGCAATGCCCTGTCCCGCC 924
QY 2377 TCATGCTGAGGCTGCTCGGGCAGCTGCTGCCCAATTAAGACTGGTCCAGACCTGTCTG 2436
DB |||||
925 TCATGCTGAGGCTGCTCGGGCAGCTGCTGCCCAATTAAGACTGGTCCAGACCTGTCTG 2436
QY 2437 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2465
DB |||||
985 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1013

RESULT 8
US-09-833-381-1293
; Sequence 1293, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1293
; LENGTH: 1155
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(1155)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1293

Query Match 38.3%; Score 946.8; DB 9; Length 1155;
Best Local Similarity 99.2%; Pred. No. 7.2e-220;
Matches 962; Conservative 0; Mismatches 7; Indels 1; Gaps 1;
QY 1477 GGCAAGGCACACCTGAAGCGTGCCATCCTGGGCGAGGAGGCGCTCGCGCTGCACGCC 1536
DB |||||
133 GGCAAGGCACACCTGAAGCGTGCCATCCTGGGCGAGGAGGCGCTCGCGCTGCACGCC 192
QY 1537 CTGTGCGCGCTCTCGCGAGGTGGACCTGTCTGGGCTGTGATCTCCACAGCGTGCAG 1596
DB |||||
193 CTGTGCGCGCTCTCGCGAGGTGGACCTGTCTGGGCTGTGATCTCCACAGCGTGCAG 252
QY 1597 CGCTCACTGGCCAGTATGCGGAGCTCGACCGTGAAGATGATCTTGTGAGGCTGCCGAG 1656
DB |||||
253 CCCCCACCGGCAAGTATGCGGAGTCTGACCGTGAAGATGATCTTGTGAGGCTGCCGAG 312
QY 1657 GCCCGGACATCCAGCCTAAGACCCACAGAGCCAGAGGCGAGATGCCAGCTGTGCC 1716
DB |||||
313 GCCCGGACATCCAGCCTAAGACCCACAGAGCCAGAGGCGAGATGCCAGCTGTGCC 372
QY 1717 CAGGGGAAGGGGCTGACATCTTCATCGGCTGGGGCCCTGTCTGTGTTCTCAGCCAG 1776
DB |||||
373 CAGGGGAAGGGGCTGACATCTTCATCGGCTGCCGGCCCTGTCTGTGTTCTCAGCCAG 432
QY 1777 AACCGGTGGCGCTGTGGGCGCCCTCCACTGACCCGAGGAGGCGGCTTTGGCTC 1836
DB |||||
433 AACCGGTGGCGCTGTGGGCGCCCTCCACTGACCCGAGGAGGCGGCTTTGGCTC 492

QY 1837 AGCTTCGGGGAGACTCGCTGTCTCTCATCGCTGCGGTCAATTCAGGAGCCAGGCGCG 1896
DB |||||
493 AGCTTCGGGGAGACTCGCTGTCTCTCATCGCTGCGGTCAATTCAGGAGCCAGGCGCG 552
QY 1897 GGGGCTGGGCTCAAGAGGAGGCGACTACATTTGTGTAGTGAATGGGAGGAGGAGGAG 1956
DB |||||
553 GGGGCTGGGCTCAAGAGGAGGCGACTACATTTGTGTAGTGAATGGGAGGAGGAGGAG 612
QY 1957 TGGAGACACCGCGAGGTGGTGAAGGCTGGAAGGCTGCGGGAGAGGCGGCGGCGAGCCTG 2016
DB |||||
613 TGGAGACACCGCGAGGTGGTGAAGGCTGGAAGGCTGCGGGAGAGGCGGCGGCGGAGCCTG 672
QY 2017 CAGGTGGTGTCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2076
DB |||||
673 CAGGTGGTGTCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 732
QY 2077 CTGCTGGGCGCCAGGCGGCTTCTTAAGAGCCAGAGGAGGAGGAGGAGGAGGAGGAG 2136
DB |||||
733 CTGCTGGGCGCCAGGCGGCTTCTTAAGAGCCAGAGGAGGAGGAGGAGGAGGAGGAGGAG 792
QY 2137 TCCAGTGGGCGCAGTCCCGGCGCTCTCTCAACTGAGCCGAAAGGCCAGGAGGAGGAG 2196
DB |||||
793 TCCAGTGGGCGCAGTCCCGGCGCTCTCTCAACTGAGCCGAAAGGCCAGGAGGAGGAGGAG 852
QY 2197 ACTGAGGCTGCCCCAGCCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2256
DB |||||
853 ACTGAGGCTGCCCCAGCCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 912
QY 2257 CACCCAGAGGTGGCGCTGAGGCGCAGGATCCCTGACGCCCTCAGGCCCTGAGCTGGCT 2316
DB |||||
913 CACCCAGAGGTGGCGCTGAGGCGCAGGATCCCTGACG-CCTCAGGCCCTGGCTCCAGCTGG 971
QY 2317 CAGCAAGCACCGAGCATGCCCTCCCGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2376
DB |||||
972 CAGCAAGCACCGAGCATGCCCTCCCGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1031
QY 2377 TCATGCTGAGGCTGCTCGGGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2436
DB |||||
1032 TCATGCTGAGGCTGCTCGGGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1091
QY 2437 AAAAAAAAAA 2446
DB |||||
1092 AAAAAAAAAA 1101

RESULT 9
US-10-112-944-558/c
; Sequence 558, Application US/10112944
; Publication No. US20040048249A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Yang, Yonghong
; APPLICANT: Weng, Gezhi
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Xue, Aidong J.
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Wehrman, Tom
; APPLICANT: Ghosh, Malabika
; APPLICANT: Wang, Dunrui
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wang, Zhiwei
; TITLE OF INVENTION: No. US20040048249A1el Nucleic Acids and
; FILE REFERENCE: 805A
; CURRENT APPLICATION NUMBER: US/10/112,944
; CURRENT FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: US 09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 09/491,404
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: US 09/496,914
; PRIOR FILING DATE: 2000-02-03

;; PRIOR APPLICATION NUMBER: US 09/515,126
;; PRIOR FILING DATE: 2000-02-28
;; PRIOR APPLICATION NUMBER: US 09/519,705
;; PRIOR FILING DATE: 2000-03-07
;; PRIOR APPLICATION NUMBER: US 09/540,217
;; PRIOR FILING DATE: 2000-03-31
;; PRIOR APPLICATION NUMBER: US 09/552,929
;; PRIOR FILING DATE: 2000-04-18
;; PRIOR APPLICATION NUMBER: US 09/577,408
;; PRIOR FILING DATE: 2000-05-18
;; NUMBER OF SEQ ID NOS: 924
;; SOFTWARE: pt_PL_genes Version 5.0
;; SEQ ID NO 558
;; LENGTH: 1046
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (1)..(1046)
;; OTHER INFORMATION: n = a,t,c or g
US-10-112-944-558

Query Match 27.3%; Score 674; DB 13; Length 1046;
Best Local Similarity 97.9%; Pred. No. 1.1e-153;
Matches 683; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1740 CCAATGGCTGGGCCCCCTGCTGTGTTCTCAGCCAGAACCGGTGGCGGCTGTGGGGCC 1799
DB |||||
702 CCCTGCCAGGGGCCCTGTCTGTGTTCTCAGCCAGAACCGGTGGCGGCTGTGGGGCC 643

QY 1800 CGTCCACTCACCGGAGGAGGGCGCTTTGGCTCAAGCTTCGGGGAGACTGCGCTGT 1859
DB |||||
642 CGTCCACTCACCGGAGGAGGGCGCTTTGGCTCAAGCTTCGGGGAGACTGCGCTGT 583

QY 1860 CCTCATCGCTGCCCTCAATTCAGGGAGCCAGGGCGCGCGCTGGCTGAAGAGGGCGGA 1919
DB |||||
582 CCTCATCGCTGCCCTCAATTCAGGGAGCCAGGGCGCGCGCTGGCTGAAGAGGGCGGA 523

QY 1920 CTAATTGTGTAGTGAATGGGAGCCATGAGTGTGTGAGACACCGGAGGTGTGAC 1979
DB |||||
522 CTAATTGTGTAGTGAATGGGAGCCATGAGTGTGTGAGACACCGGAGGTGTGAC 463

QY 1980 GGAGCTCAAGCTCGGGAGAGGGCGCGCGAGCTTCAGTGTGTGCTGTGCTGCCAG 2039
DB |||||
462 GGAGCTCAAGCTCGGGAGAGGGCGCGAGCTTCAGTGTGTGCTGTGCTGCCAG 403

QY 2040 CTCTAGACTGCCCCAGCTTGGGGGACCGCCGCGCGCTCTCTGCTGGGCCCCAGGGGGCTTCT 2099
DB |||||
402 CTCTAGACTGCCCCAGCTTGGGGGACCGCCGCGCGCTCTCTGCTGGGCCCCAGGGGGCTTCT 343

QY 2100 AAGGAGCCAGAGGAGCATGTTTCAGACCCCGGATCCACGTGGGCGAGTCCCGGCC 2159
DB |||||
342 AAGGAGCCAGAGGAGCATGTTTCAGACCCCGGATCCACGTGGGCGAGTCCCGGCC 283

QY 2160 CCTCTCAACTGGAGCCGAAAGGCCAGAGGCAAGACTGGAGGCTGCCCGCCAGCCCTG 2219
DB |||||
282 CCTCTCAACTGGAGCCGAAAGGCCAGAGGCAAGACTGGAGGCTGCCCGCCAGCCCTG 223

QY 2220 TGCCCCAGTGAAGCAGCTCCGCTCATCTTGAAGACCCAGGGTGGCGGTGAGGGCC 2279
DB |||||
222 TGCCCCAGTGAAGCAGCTCCGCTCATCTTGAAGACCCAGGGTGGCGGTGAGGGCC 163

QY 2280 AGGATCCCTGCAGCCCTCAGCCTGGCTCCAGCTGGCAGCAAGCAGCATGCCCTC 2339
DB |||||
162 AGGATCCCTGCAGCCCTCAGCCTGGCTCCAGCTGGCAGCAAGCAGCATGCCCTC 103

QY 2340 CCCACCCAGAGACCTCCGGGCAATGCTGTCCCGCTCATGCTGGAGGCTCCCTCGGGC 2399
DB |||||
102 CCCACCCAGAGACCTCCGGGCAATGCTGTCCCGCTCATGCTGGAGGCTCCCTCGGGC 43

QY 2400 ACCTGCCTGCCATTAAAGACTGGTTCAGACTGTCTGA 2437
DB |||||
42 ACCTGCCTGCCATTAAAGACTGGTTCAGACTGTCTGA 5

RESULT 10
US-10-239-607-52
;; Sequence 52, Application US/10239607
;; Publication No. US20030219761A1
;; GENERAL INFORMATION:
;; APPLICANT: Saatchioglul, Fahri
;; TITLE OF INVENTION: No. US20030219761A1 Prostate-Specific or
;; TITLE OF INVENTION: Testis-Specific Nucleic Acid Molecules, Polypeptides, and
;; TITLE OF INVENTION: Diagnostic and Therapeutic Methods
;; FILE REFERENCE: 50218/003002
;; CURRENT APPLICATION NUMBER: US/10/239,607
;; CURRENT FILING DATE: 2002-09-24
;; PRIOR APPLICATION NUMBER: PCT/US01/09410
;; PRIOR FILING DATE: 2001-03-23
;; PRIOR APPLICATION NUMBER: US 60/191,929
;; PRIOR FILING DATE: 2000-03-24
;; NUMBER OF SEQ ID NOS: 73
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 52
;; LENGTH: 2100
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-10-239-607-52

Query Match 20.1%; Score 496.4; DB 16; Length 2100;
Best Local Similarity 56.8%; Pred. No. 1.5e-110;
Matches 977; Conservative 0; Mismatches 731; Indels 12; Gaps 3;

QY 321 GGGCTGTGATCCCTGACCGCAGATCCAGTGCGCCAGCTCCAGAGCGCAGGGCCAGAT 380
DB |||||
69 GGGCTGTGATCCCTTGCACAAACCGCGCGAGTAAATTCAGAAATCAAGAGAGCTGCTTT 128

QY 381 TCACAGCAGATTGACAGAGGCTGCAGATGCGGACGGCGCTGAGAACCTTACAGAGC 440
DB |||||
129 GAATCAGCAGATCTGAAAGCCGTCGGATGAGAGCCGAGCGGAGAAACCTTCTGAAGT 188

QY 441 CACCAGCAACACCGGGTGAGAGACGGTTCGCCCTGGAGCTGAGCTACGCTCAACTCAA 500
DB |||||
189 GGCACAACTCAAGGTGGGGAGCAGTGGCGCTGGAGCTGAGCTTCGCTCAACTCAGA 248

QY 501 CCTGCAGCTCTGAAGAGAGGAGCTGGAGAGCTCAGCGGTGGCGTGGAGCCCTGGCCGCGCA 560
DB |||||
249 CCTGCAGTCTCAAGGAGAGCTGGAGGCGCTGAAACATCTCGGTGGCGCTTATCAGAA 308

QY 561 TGGAGGAGAGCTGTCTACTTCCCTCATGATCCCTGGCGCTGAGGAGACCAAGGAGCT 620
DB |||||
309 CACAGAGGAGGCGATTACAGATTCCCTGATTCCTCTTGGCGCTGAGGAGAAAGAGAGCT 368

QY 621 GGACTGCTTACACCGCTGAAGGAGCTGATCTCAGTGCATCTTTGGAGAGGACGGCGCTC 680
DB |||||
369 CGACTTTGCGAGTCTCTCAGGATTTATCTCGAACATTTACAGTGAAGATGCTATTT 428

QY 681 CTAGAGGCGAGAAATCAGGAGCTGGAGGCGCTTCGGCAGGCCATGCGGACCCCGCG 740
DB |||||
429 ATATGAAGATGAATTCAGATCTTATGATCTGAGACAAAGCTTGTGCGAGCGCTAGCG 488

QY 741 GAATGAGTGGCGCTGGAGCTGCTCAGAGCTATTACAAACAGCTGTGCTTCTGGATGC 800
DB |||||
489 GGATGAGGCGGGGTGGAATCTGTGATGACATCTTATCAGCTGGCTTGTTCGAGAG 548

QY 801 GCGCTTCTCACCCTCCAGAGGCTTCGGGCTCTTTCACCTGGTACGATCGCTCGCTTAC 860
DB |||||
549 TCGATTCTTCCCGCCACACGGCAGATGGGACTCTCTTACCTGGTATGACTCTCTCAC 608

QY 861 TGGGGTCCCGGCCAGCAGCGTGCCTGGCCCTTCAGAGGCGAGCTTCTTTCACAT 920
DB |||||
609 TGGGGTCCCGGTTCAGCCAGCAGAAACCTGCTGCTGGAGAGGCCAGTGTCTGTTCACAC 668

QY 921 CGGTGCGCTCCACAGCAGATTGGGCGCGCCAGAGCCGCTCTCTGACCCGAGGGTGC 980
DB |||||
669 TGGGGCGCTCTACACCCAGATTGGGACCCCGGTGCGATCGGACGAGCGGCTGGGCTGGA 728

QY	981	CGCGCTATGGAGGCTTTCAGAGGGCCGCTTGGGGCCCTTACGCTCTCTGAGGGAGAACTT	1040
DB	729	GAGTGCCATAGATGCTTTTCAGAGAGCCGACGGGTTTTAAATTTACCTGAAAGACACATT	788
QY	1041	CTCCCATGGCCGAGCCACAGACATGAGCGCTGCGTCCCTCTTGCGCATGTGGAGCAGCTCAT	1100
DB	789	TACCCATATCTCAAAGTTACGACATGAGCCCTGCCATGCTCAGCGTGTCTCGTCAAAATGAT	848
QY	1101	GATGGCCCGAGCCCGAGGAATGTGTGTTTGAGGGCTCTCACCACTGCCCTCCATGGCCCC	1160
DB	849	GCTTGCACAGCCCAAGAAACGCTGTTGAGAAATCAGCTTCCCTG-----GGATCCG	902
QY	1161	CCAAAGACTGCTGGCCCGCAGCTGCGCTTGGCGCAGAGGCGCCCGAGGTGGCAGCCGAGTA	1220
DB	903	GAATGAATTTCTTCATGCTGTGTGAAGTGGCTCAGAGGCTGCTAAGCTGGGAGAGTCTA	962
QY	1221	CAGCTAGTGACCGGACCATGGGCCCAAGCCACCCGTCACAGTACGTGCTGTCTCTG	1280
DB	963	CCAAACAGCTACACGACGCCATGAGCCAGGCGCGGTGAAAGAGAACTCCCTACTCTGTC	1022
QY	1281	GACTGCCCTGTGTGATGCTCAAGGCGGAGTACTTCCGCTCCCTGGTCCCACTTACCACGTAGC	1340
DB	1023	GGCCAGCTTAGCTGTGTGTAAGGCCCAACCACTACGCGCCCTGGCCCACTACTTCACTGC	1082
QY	1341	CATGGCCCTCTGGACCGGCTCCCAGCGACCGGAGGAGAGCTCCCCACGACACGACGAGGT	1400
DB	1083	CATCTCTCTCATCGACCACACAGGTGAAGCGCAGGACCGGATCTGGACCAACAGAGAGTG	1142
QY	1401	CTTCTCTGAGC----CCCCCACTCTCTTAAGCCCCGAGGCCCTGTGTC--TGCGCGAGGA	1454
DB	1143	CCTGTCGACGCTTACGACCACATGCCAGAGGGGCTGACACCTTGGCCACACTGAAGAA	1202
QY	1455	GCTGGAGAGCGCAGGACGCTTGCCAGGACACACTGAAGCGTGCCATCTCTGGGGCAGGA	1514
DB	1203	TGATCAGACGGCCGACAGCTTGGGGAAGTCCCACTTGGCAGAGCCATGGCTCATACGA	1262
QY	1515	GGAGGCGCTGCGGTGCAACGCCCTGTGTCGCGCTCTGCGCGAGTGGACCTGCTTCGGGC	1574
DB	1263	GGAGTCGTGCGGAGGCGAGCCTCTGCAAGAGCTGCGGAGCATTTGAGTGTCTACAGAA	1322
QY	1575	TGTGATCTCCAGACGCTGCAGCGCTCACTGGGCCAAAGTATGCGGAGCTGGACCGTGAGGA	1634
DB	1323	GGTGCTGTGTGCCGACACAGGAACGCTCCCGGCTCACGTAGCCCGACCCAGGAGGAGGA	1382
QY	1635	TGACTTCTGTGAGGCTGCCGAGGCCCGGACATCCAGCCTTAAGACCCACAGAGCCAGA	1694
DB	1383	TGACCTGCTGAACCTGATCGACGCCCCAGTGTGTTGCTTAAACTGAGCAAGAGTTGA	1442
QY	1695	GGCCAGGATGCCACGCTGTCCCAGGGGAGGGGCCCTGACATCTCCATCGGCTGGGGCC	1754
DB	1443	CATTATATTGCCCGAGTTCTCCAAGCTGACAGTCACGGACTTCTTCAGAGCTGGGCC	1502
QY	1755	CCTGTCTGTGTTCTCAGCCAAAGAACGGTGGCGGCTGGTGGGGCCCGTCCACCTGACCCG	1814
DB	1503	CTTATCTGTGTTTTTGGCTTAAAGCGGTGGAAGCGCTCTCGAAGCATCCGCTTCACTGC	1562
QY	1815	AGGAGGGGCGGTTTTGGCTTCACGCTTGGGGAGACTCGCCTGTCTCTCATCGCTGCCGT	1874
DB	1563	AGAAGAAGGGAACTTGGGGTTTCACTTTGAGAGGGAACGCCCGCTTCAGGTTCACTTCT	1622
QY	1875	CATTTCAGGAGCAGGCGCGCGGCTGGCCCTGAAGGAGGGCGACTACATTGTGTCACT	1934
DB	1623	GGATCCTTACTGCTCTGCGCTCGTGGCAGGAGCCCGGGAAGAGATTATTGTCTCCAT	1682
QY	1935	GAATGGGACGCCATGCAAGTGTGTGGAGACACCGGAGGTGGTGAACGAGCTGAAGGCTGC	1994
DB	1683	TCAGCTTGTGGATTGTAAGTGGCTGACGCTGAGTGAAGTTATGAAGCTCTCTGAAGCTT	1742
QY	1995	GGCAGAGGGGGCGCCAGCCTGCAGTGTGTGCTGCTG 2034	
DB	1743	TGCGGAGGACGAGATCGAGATGAAGTCTGTGAGCTCTCTG 1782	

```

RESULT 11
US-10-239-607-51
; Sequence 51, Application US/10239607
; Publication No. US20030219761A1
; GENERAL INFORMATION:
; APPLICANT: Saatchioglou, Fahri
; TITLE OF INVENTION: Testis-Specific Nucleic Acid Molecules, Polypeptides, and
; TITLE OF INVENTION: Diagnostic and Therapeutic Methods
; FILE REFERENCE: 50218/003002
; CURRENT APPLICATION NUMBER: US/10/239,607
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: PCT/US01/09410
; PRIOR FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 60/191,929
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 51
; LENGTH: 3526
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-239-607-51

```

```
Query Match      20.1%; Score 496.4; DB 16; Length 3526;
Best Local Similarity 56.8%; Pred. No. 1.7e-110;
Matches 977; Conservative 0; Mismatches 731; Indels 12; Gaps 3
```


Matches	976;	Conservative	0;	Mismatches	732;	Indels	12;	Gaps	3;
QY	321	GGGCTGTGACTCCCTGACGCGAGATCCAGTGGCGGCAGCTGCAGAGCGCGAGGGCCAGAT	380						
Db	79	GGGCTGTAAATCCCTTGCACAAACCGCGCGAGTAAATTGCAGAAATCAAGAGCTGCTTT	138						
QY	381	TCACGACGAGATTGACAGGAGCTGCAGATCGGACGGCGCTGAGAACCTTCTACAGAGC	440						
Db	139	GAATCAGCAGATCTCTGAAGCCGTGCGGATGAGGACCGGACCGGAAACCTTCTGAAAGT	198						
QY	441	CACCAGCAACAACCGGCTGAGAGACCGTGCCTTGAGGTGAGCTACGTCAACTCAA	500						
Db	199	GGCCACAAACTCAAGGTGCGGAGCAAGTCGCGCTGGAGCTGAGCTTCGTCAACTCAGA	258						
QY	501	CCTGCAGCTGCTGAAGGAGGAGCTGGAGCAGCTCAGCGGTGGCGTGGACCTTGGCGCGCA	560						
Db	259	CCTGCAGATGCTCAAGGAAAGCTGGAGGGGCTGAACATCTCGGTGGGCGTCTATACGAA	318						
QY	561	TGGGAGCGAAGGTGTCACTGTCGCCATGATCCCCCTGGCGCTGAAGGAGACCAAGGAGCT	620						
Db	319	CACAGAGGAGCAATTTACGATTCCTCTGATTCCTCTGGCTGAAGGAAACGAAAGACGT	378						
QY	621	GGACTGGTCTACCGCTGAGGAGCTGATCTCAGTGCACTTTGAGAGAGCGCGCTC	680						
Db	379	CGACTTTGAGTCGTCTCAAGGATTTATCCTCGAAATTTACACTGAGATGGCTATTT	438						
QY	681	CTACGAGCAGAAATCAGGAGCTGGAGGCGCTCGCGCAGGCCATCGGACCCCGACGCG	740						
Db	439	ATAAGAAGTGAATTCGAGATCTTATGGATCTGAGACAGCTTGTGCGACGCTACCGC	498						
QY	741	GAATGAGTGGGCTCGAGCTGTCTACAGGCTATTACCAACAGCTGTGCTTCCTGGATGC	800						
Db	499	GGATGAGCGCGGGTGGAACTGCTGATGACATACTTTCATCCAGCTGGCTTTGTGAGAG	558						
QY	801	CGCGTCTCTCACCCCTGCCAGGACCTCGGGCTCTTCTTCCATGGTACGACTGCGTTAC	860						
Db	559	TCGATTTCTTCCGCGCCACACGCGAGATGGGACTCCTGTTCACTGGTATGACTCTCTCAC	618						
QY	861	TGGGGTCCCGGCCCAGCAGCGTGCCCTGGCTTCGAGAAGGGCAGCGTCTCTTCAACAT	920						
Db	619	CGGGTTCGGTCAAGCAGCAGAACCTTGCTGCTGGAGAGGCCAGTGCTCTGTTTCAACAC	678						
QY	921	CGGTGCCCTTCACACGAGATTGGGGCGCCAGGACCGCTCTCTGCACCGAGGGTGCCCG	980						
Db	679	TGGGGCCCTCTACACCCAGATTGGGACCCGGTGCATCGCGACGCGAGCTGGGCTGGA	738						
QY	981	CCGCGCTATGGAGCCTTCCAGAGGGCGCGCTGGGGCCTTCAGCCTCTGAGGAGAGACTT	1040						
Db	739	GAGTGCCCATAGATGCCCTTCAGAGAGCCGACGGGGTTTAAATTTACCTGAAGACACAT	798						
QY	1041	CTCCATGCGCCGACGCCAGACATGAGCGTGCCTGCTCTGCGCATGGAGCAGAGCTCAT	1100						
Db	799	TACCCATCTCCAAGTTACGCATGAGCCTGCGCATGCTCAGCGTGTCTGCAAAATGAT	858						
QY	1101	GATGGCCACGGCCGAGGAATGTGTGTTGAGGCGCTCTCAACCTTCCTCCATGCGCCC	1160						
Db	859	GCTTGCAAGCCCAAGAAACGGTGTGAGAAAAATCAGCCTTCTCTG-----GGATCCG	912						
QY	1161	CCAAGACTGCTGGCCAGCTGCGCTGGCGGAGGAGCGCCAGGTGGAGCGCCAGTA	1220						
Db	913	GAATGAATTTCTCATGCTGTGAAGTGGCTTCAGAGGGCTGTCAAGGTGGAGAGGTCTA	972						
QY	1221	CAGGCTAGTGACCCGACCATGGCCACCGCCAGCCGCTCCAGACTAGTGTCTCTCTCTG	1280						
Db	973	CCAACGACTACCGACCCATGAGCCAGCGCGCGGTGAAGAGAAATCCCTTACTCTCTG	1032						
QY	1281	GACTGCCCTTGATGTCAAGGCCGAGTACTTTCCGCTCCCTGGCCCACTTACCACGTAGC	1340						
Db	1033	GGCCAGCTTAGCTGTGAAGCCCACTACTAGCGGCCCTGGCCCACTACTTCACTGC	1092						
QY	1341	CATGGCCCTCTGCACGGCTCCCGACGACCGAGGGAGAGCTCCCAAGCAGAGCAGGT	1400						
Db	1093	CATCTCTCTCATCGACACACAGGTGAAGCCAGGACCGGATCTGGACCAACAGGAGGTG	1152						

QY	1401	CTTCTGTGAGC-----CGCCCACTCTCTTAAGCCCCGAGGCCCTGTGTC--TGCCGACGGA	1454
DB	1153	CCTGTCCCAAGCTCTACGACCAACATATGCCAGAGGGGCTGACACCCCTTGCCACACTGAAGAA	1212
QY	1455	GCTGGAGAGCGCAGGAGCTTGCGAAGGCACACCTGAAGGTGCCATCTGGGGGACGGA	1514
DB	1213	TGATCAGACGCGCCGACAGCTTGGGGAAGTCCCACCTTGCGCAGAGCCATGGCTCATCACGA	1272
QY	1515	GGAGCGCTGCGGTGCAACGCCCTGTGCGCGCTCTCGCGCAGGTGACCTGTCTGGGGC	1574
DB	1273	GGAGTCTGTCGGGAGGCAAGCCCTCTCAAGAAGCTGCGGAGCATTGAGTGCTACAGAA	1332
QY	1575	TGTGATCTCCACAGACGTGTGACGGCTCACTGCGCCAGTATGCGGAGCTGCACCGTGAGGA	1634
DB	1333	GGTGCTGTGTGCGCA CAGGAAGCTCCCCGGCTCACGTACGCCGACCCAGAGGAGGA	1392
QY	1635	TGACTTCTGTGAGGCTGCCGAGGCCCGGACATCCAGCCTTAAGACCCACCAGAAAGCCAGA	1694
DB	1393	TGACCTGCTGAACCTGATCGAGGCCCCAGTGTTGTGTAAACTGTGACGACAGAGTTGA	1452
QY	1695	GGCCAGAGATGCCACGCCTGTCTCCAGGGGAAGGGCCCTTGACATCTTCAATCGGCTGGGGCC	1754
DB	1453	CATTATATTGCCCCAGTCTCTCCAAGCTGACAGTCAOGGACTTCTTCAGAAAGCTGGGCC	1512
QY	1755	CCTGCTGTGTTCTTCAGCCAAAGAACCGTGGCGGCTGGTGGGCCCGCTCCACCTGACCCG	1814
DB	1513	CTTATCTGTGTTTTCGGCTAAACAGCGGTGACGCTCTCTGAAGCATCCGCTTCACTGC	1572
QY	1815	AGGAGAGGCGGCTTTGGCCCTCA CGCTTGGGGAGACTCGCCTGTCTCTCATCGCTGCCGT	1874
DB	1573	AGAAGAGGGGCACTTGGGGTTCACTTGAGAGGAACGCCCGCTTCAGGTTCACTTCTCT	1632
QY	1875	CATTCCAGGAGGACGACGCGCGCGGCTGGCTGAAGAGGGGCGACTACATTGTGTCAGT	1934
DB	1633	GGATCTCTTA CTGCTCTGCTCGGTGGCAGGAGCCCGGAAGGAGATTATATTGTCTCCAT	1692
QY	1935	GAATGGGCACCCATGCAGTGTGTGAGACAACGCGGAGGTGGTGACGAGCTGAAGGCTGC	1994
DB	1693	TCAGCTGTGTGATTGAATGGCTGACGCTGAGTGAGGTTATGAAGCTGCTGAAGAGCTT	1752
QY	1995	GGGAGAGCGGGCGCCACGCCTGCAAGTGTGTGCTGTGCTG 2034	
DB	1753	TGCGAGAGCCAGATCGAGATGAAGTCTGTGAGCCCTCTG 1792	

```

RESULT 13
US-10-108-260A-2397
; Sequence 2397, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: NO. US20040005560A1el full length cDNA
; FILE REFERENCE: HI-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2397
; LENGTH: 3500
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-108-260A-2397

```

Query Match	20.0%;	Score 494.8;	DB 16;	Length 3500;
Best Local Similarity	56.7%;	Pred. No. 4.1e-110;		

QY 321 GGGCTGTGACTCCCTCAGCGCATGCCAGTGCAGAGCCCAGAGGCCAGAT 380
| | | | | | | | | | | | | | | |
Db 105 GGGCTGTAATCCCTTCACAAACCGCCGAGTAAATTGCAGATCAAGAGTGCTTT 164

QY 381 TCACCAGCATTTGAAGAAGTGTCAGATGCGGACGGCGCTGAGAACCTCTCAGAGC 440

1239 TGATCAGCAGCGCCGACAGCTGGGGAAGTCCCACTTTGCGCAGAGCCATGCTCATCAGA 1298
1515 GGAGGCGCTGGCGCTGCAAGCGCTGTGCGCGCTGCGAGCTGACCTGCTCGGGC 1574
1299 GGAGTGGTGGGAGGCGAGCCTCTGCAAGAAGCTGCGGAGCATTTGAGGTGTTACAGAA 1358
1575 TGTGATCTCTCCAGACGCTGAGCGCTCACTGGGCCAAGTATGCGGAGCTCGACCGTAGGA 1634
1359 GGTGCTGTGTGCGCACAGCAAGCTCCCGGCTCAGTACGCCAGCACCCAGGAGGAGGA 1418
1635 TGACTTCTGTGAGGCTGCGAGGCGCCCGGACATCCAGCCTTAAGACCCACAGAGCCAGA 1694
1419 TGACCTGCTGAACCTGATCGACGCCGCCAGTGTGTTGCTTAAACTGAGCAAGAGGTGA 1478
1695 GGCAGAGTACCCAGCCTGTCCAGGGGAGGGGCGCTGACATCTTCCATCGGCTGGGGCC 1754
1479 CATTATATTGCCCAAGTTCTCCAAGCTGACAGTACGAGCTTTCTTCCAGAGCTGGGCC 1538
1755 CTTGCTGTGTTCTCAGCCAAAGAACCGGTGGCGGCTGTGGGGCCCGTCCACCTGACCCG 1814
1539 CTTATCTGTGTTTTCGCTAAAGCGGTGGACGCTCTCGAAGCATCCGCTTCACTGC 1598
1815 AGAGAGGCGGCTTTGGCCTCAGGCTTCGGGAGACTCGCCTGTCTCATCGCTGCGCT 1874
1599 AGAAGAGGGGACTTGGGGTTTCACTTTGAGAGGGAACGCCCCGTTTCAAGTTCATCTTCT 1658
1875 CATTCCAGGAGGAGCGCGCGCTGGCTGAAGGAGGGGAGCTATCATTTGTGTGCTG 1934
1659 GATTCCTTACTGCTCTGCTGGTGGCAGAGCGCGGAGGAGATTTATTTGCTTCCAT 1718
1935 GAATGGGCGACCATGACAGTGGTGGACACCGGGAGGTGGTGCAGGCTGAAGGCTGC 1994
1719 TCAGCTTGTGATTTGTAAGTGGCTGACGCTGAGTGAAGTTATGAAGCTGCTGAAGAGCTT 1778
1995 GGAGAGGCGCGCGCGCTGAGCTGAGGTTGGTGTGCTGCTGCTG 2034
1779 TGGCAGGAGCAGATCGAGATGAAGTCGTGAGCCTCTG 1818

RESULT 14

US-09-895-040A-2
; Sequence 2, Application US/09895040A
; Patent No. US2002012347A1
; GENERAL INFORMATION:
; APPLICANT: Shannon, Mark
; TITLE OF INVENTION: HUMAN GTP-RHO BINDING PROTEIN 2
; FILE REFERENCE: ABOMICA-11
; CURRENT APPLICATION NUMBER: US/09/895,040A
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 09/864,761
; PRIOR FILING DATE: 2001-05-23
; NUMBER OF SEQ ID NOS: 180
; SOFTWARE: Aesomica Sequence Listing Engine
; SEQ ID NO 2
; LENGTH: 2061
; TYPE: DNA

165 GAATCAGCAGATCTGAAAGCCATGGGATGAGACCGGAGCGGAAACCTTCTGAAGT 224
441 CACAGCAACAACCGGTGAGAGAGCGGTGCGCTGGAGCTAGCTACGTCAACTCAA 500
225 GGCACAAACTCAAGGTGCGGAGCAAGTGGCGCTGGAGCTGAGCTTCTCAACTCAGA 284
501 CTTGAGCTCTGAAGGAGAGCTGAGAGGAGCTCAGCGGTGGGTGAGACCCCTGGCCGCA 560
285 CTTGCGATCTCAAGGAAGAGCTGGAGGGCTGAACATCTCGGTGGGCTCTTATCAGAA 344
561 TGGGAGCGAGCTGCTCACTCTCCCATGATCCCTCGGCTGAAGGAGACCAAGAGAGCT 620
345 CACAGAGGAGGCAATTACGATTTCCCTGATTTCTTGGCTGAGAGAAACGAAAGAGCT 404
621 GGAATGCTGTACACCGCTGAAGGAGCTGATCTCAGTGCACATTTGGAGAGACCGGCGCT 680
405 CGACTTTGCGAGTCTGCTCAAGGATTTTATCTTGGAAACATTACAGTGAAGATGGCTATT 464
681 CTAGGAGGAGAAATCAGGAGCTGGAGGCGCTCGGCGAGGCCATCGGACCCCGCCAGCG 740
465 ATATGAAGATGAATTCAGATCTTATGGAATCTGAGACAGACTTGTTCGGAGCGCTAGCCG 524
741 GAATGAGTGGGCTGAGCTGCTCAGAGCTTATTAACAACAGCTGTGCTTCTTGGATGC 800
525 GGAATGAGCGGGTGAATCTGTGATGACATCTTCACTGAGCTGGCTTTGTGAGAG 584
801 GCGCTTCTTCACTCCCTGCGAGGAGCTCGGCTCTTCTTCACTGAGTACAGCTGCTTAC 860
585 TCGATTTCTCCGCGCCACACGAGATGGGACTCTCTTCACTGTTATGACTCTCTCAC 644
861 TGGGTTCCGCGCCAGCAGCTGCGCTTCCGAGGAGGAGCTTCTTCAACAT 920
645 CGGGGTTCCGCTCAGCCAGCAGAACCTGCTGCTGGAGAGGCCAGTGTCTGTTTCAACAC 704
921 CGGTGCTCTCACACGAGATTTGGGCGCGCCAGGAGCGCTCTCTGACCGAGGCTGCGCG 980
705 TGGGCGCTCTACACCCAGATTGGACCCGGTGTGATCGGACGCGAGCTGGGCTGGA 764
981 CCGCGCTATGAGGCGCTTCCAGAGGCGCTGGGCGCTTTCAGCTCTGAGGAGAACTT 1040
765 GAGTGCCATAGATGCTTTTCCAGAGCGCGAGGGGTTTAAATTTACCTGAAAGACACATT 824
1041 CTCCATGCGCGAGCCAGACATGAGCGCTGCTGCTGCTGCGACCTGGAGAGCTCAT 1100
825 TACCATATCTCAAGTTACGACATGAGCGCTGCGCATGCTGAGCTGCTGCTCAAAATGAT 884
1101 GATGCGCCAGGCCAGGAATGTGTTTGGGGCTCTTCAACACCTGCTCCATGGGCC 1160
885 GCTTGCACAGCCCAAGAACGCTGTTGAGAAATCAGCCTTCTCTG-----GGATCCG 938
1161 CCAAGACTGCTGGCGCAGCTGCGCTGGCGCAGAGGCGCCCGAGGTGGCAGCGAGTA 1220
939 GAATGAATTTCTATGCTGTGTAAGGTGGCTCAGAGGCTCTAAGGTGGAGAGGTCTA 998
1221 CAGCTAGTACCGGACCAATGCGCCAGCCAGCCCGTCCAGACTACGCTGCTCTCTG 1280
999 CCAACAGCTACCGCAGCCATGAGCGCGCGGTGAAAGAAACATCCCTACTCTCTG 1058
1281 GACTGCTGCTGATGATCAAGGCGGAGTACTTCCGCTTCCCTGGCCCACTACCAAGTAGC 1340
1059 GGCAGACTTAGCTGCTGTAAGGCCACCACTAGCGGCGCTGGCCCACTACTTCACTG 1118
1341 CATGGCCCTCTGCGACGCTCCCGACGACGAGGAGAGTCCCCACGAGGAGAGGT 1400
1119 CATCTCTCTATCGACACCAAGGTGAACCGAGGACGAGTCTGACACACGAGGAGAGTG 1178
1401 CTTCTGCGAGC-----CCGCCACTCTCTTAAAGCCCGAGGCGCTGTGTC--TGCGCAGGA 1454
1179 CTTGTCAGCTCTAGACCAACATGCGAGGGGCTGACACCTTGGCCACACTGAAGAA 1238
1455 GCTGAGAGGCGCAGGAGCTTGGCAAGGACACCTGAAGGCTGCTCTGCGGCGAGGA 1514


```
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 09/864,761
; PRIOR FILING DATE: 2001-05-23
; NUMBER OF SEQ ID NOS: 180
; SOFTWARE: Acomica Sequence Listing Engine
; SEQ ID NO 1
; LENGTH: 3484
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-895-040A-1

Query Match      20.0%; Score 493.2; DB 9; Length 3484;
Best Local Similarity 56.7%; Pred. No. 9.9e-110;
Matches 975; Conservative 0; Mismatches 733; Indels 12; Gaps 3;

QY 321 GGGCTGTGACTCCCTGACGAGATCCAGTGGCGCAGCTGCAGAGCGCGAGGCCAGAT 380
Db      |||||
QY 89 GGGCTGTGAAATCCCTTGCACAAACCGCGCGAGTAAATTCAGAAATCAAAGAGCTGCTT 148
Db      |||||
QY 381 TCACACGACAGATTGACAAAGAGCTGCAGATCGGACGGGGCTGAGAACTCTACAGAGC 440
Db      |||||
QY 149 GAATCAGCAGATCTTGAAGCCGTGCGGATGAGGATCGGAGCGGAAACCTTCTGAAAGT 208
Db      |||||
QY 441 CACACGACAAACCGGCTGAGAGAGACGCTGCCCTCGAGTGAAGTACGTTCAACTCAA 500
Db      |||||
QY 209 GGGCACAACTCAAAGCTGCGGAGCAAGTGGCGCTGAGCTGAGCTTCTCAACTCAGA 268
Db      |||||
QY 501 CTTGACCTGCTGAGGAGGCTGAGGAGCTCAGCGTGGCTGAGCCCTGGCGGCA 560
Db      |||||
QY 269 CTTGACATGCTCAAAGAGAGCTGGAGGGGCTGAACATCTCGGTGGGCGTCTATCAGAA 328
Db      |||||
QY 561 TGGAGGGAAGCTGCTCACTGTCCTCCATGATCCCCCTGGGCTGAAGAGACCAAGAGCT 620
Db      |||||
QY 329 CACAGAGGAGCAATTACGATTCCTCTGATTCCTTGGCTGAAGGAAACGAAAGAGCT 388
Db      |||||
QY 621 GGACTGTCTACACCGCTGAAGAGAGTATCTCAGTGACATTTTGGAGAGAGCGGCTC 680
Db      |||||
QY 389 CGACTTTGCACTGCTCTCAAGGATTTTATCTGGAACATTTACAGTGAAGATGCTATTT 448
Db      |||||
QY 681 CTACGAGGACAAATCAGGAGCTGAGGCGCTCGGCGAGCCCATCGGACCCCGAGCG 740
Db      |||||
QY 449 ATATGAAGATGAATTCAGATCTTATGGATCTGAGACAGCTTGTGAGCGCTTAGCGG 508
Db      |||||
QY 741 GAATGAGTCGGGCTGAGCTGCTCACAGCCTATTACAACCAAGCTGTCTTCTTGGATGC 800
Db      |||||
QY 509 GGATGAGGCGGGGTGGAATCTGCTGATGACATCTTATCCAGCTGGGCTTTGTGAGAG 568
Db      |||||
QY 801 GCGCTTCTACCCCTGCGAGAGCTCGGCTCTTCTTCACTGGTACGACTCGCTTAC 860
Db      |||||
QY 569 TCGATTTCTTCCGCGCCACACGCGCAGATGGGACTCTCTGTACCTGGTATGACTCTCTAC 628
Db      |||||
QY 861 TGGGCTCCCGCCAGCAGCTGCGCTTCCGAGAGGGCAGCGTCTCTTCAACAT 920
Db      |||||
QY 629 CGGGGTTCCGGTACCCAGCAGAACCTGCTGCTGAGAGAGCCAGTGTCTTCTTCAACAC 688
Db      |||||
QY 921 CGTGTCCCTTCCACACGAGATTGGGGCGCGCCAGACCGCTCTCTGCAACCGAGGGTGC 980
Db      |||||
QY 689 TGGGGCCCTCTACACCCAGATTGGGACCCGCTGCGATCGGAGAGCGAGGCTGGGCTGA 748
Db      |||||
QY 981 CCGCGCTATGAGGCTTCCAGAGGGCGCTGGGCGCTTACGCTCTTGGAGGAGAACTT 1040
Db      |||||
QY 749 GAGTGCCTATAGATGCTTTCAGAGAGCGCGAGGGGTTTAAATTAACCTGAAAGACACAT 808
Db      |||||
QY 1041 CTCCTCATGCGCGAGCCAGACATGAGCGCTGCGTCCCTCTGCGCACTGGAGCAGCTCAT 1100
Db      |||||
QY 809 TACCCATACTCCAAGTTACGACATGAGCCCTGCCATGCTCAGCGTGTCTCGTCAAAATGAT 868
Db      |||||
QY 1101 GATGGCCCGAGCCCAAGGAATGTGTGTTGAGGGCCCTCTCACCACTGCTCTCATGGCCCC 1160
Db      |||||
```

```
869 GCTTGACAAAGCCCAAGAAAGCGTGTGTGAGAAAAATCAGCCTTCTCTG- - - - -GSAATCCG 922
1161 CCAAGACTGCTGGCCAGCTGCGCTGGCGCAGAGGAGCGCCAGGTGCGAGCCGAGTA 1220
923 GAATGAATTTCTCATGCTGGTGAAGTGGCTCAGGAGGCTGCTAAGTGTGGAGAGTCTA 982
1221 CAGGCTAGTGCACCGGACCATGCGCCAGCCAGCCCGCTCCACGACTACGTGCTCTCTCT 1280
983 CCAACAGCTACACGAGCCATGAGCCAGGCGCGGTGAAGAGAACATCCCTACTCTCTG 1042
1281 GACTGCGCTGGTGCAATGTAAGCCGAGTACTTCCGCTCCCTGGCCCACTACACAGTAGC 1340
1043 GCGCAGCTTAGCCTGTGAGGAGCCACCACTAGCGGGCCCTGGCCCACTACTTCACTGC 1102
1341 CATGGCCCTCTGCGAGCGCTCCCGAGCGACCGAGGAGAGCTCCCGACGACAGAGAGGT 1400
1103 CATCTCTCTCATCGACACACAGGTGAAGCCAGGACCGGATCTGACACACAGAGAGAGTG 1162
1401 CTTCTCTGAGC- - - - -CCCCCACCCTCTTAAGCCCGAGGCGCTGTGC- - - - -TGCGGAGGA 1454
1163 CCGTGTCCCACTCTACGACCACATGCGAGAGGGCTGACACCCCTTGGCCACACTGAAGAA 1222
1455 GCTGAGAGCGCAGGAGCTTGGCAAGGCACACCTGAAGCGTGCCTATCTGGGCGAGGA 1514
1223 TGATCAGACGCGCGCAGACTGGGGAAGTCCCACTTGCAGAGAGCCATGGCTCATACGA 1282
1515 GAGGCGCTGCGGCTGACGCGCTGTGCGCGCTCTGCGAGGTGACCTGCTTCGGGC 1574
1283 GAGTCCGTGCGGAGGCAAGCTCTGCAAGAGCTGCGGAGCATTGAGGTGCTACAGAA 1342
1575 TGTGATCTCCAGACGCTGCAGCGCTCACTGGCCCAAGTATTCGGAGTTCGACCGTAGGA 1634
1343 GGTGCTGTGTGCGCACAGGAACGCTCCCGCTCACGTACGCCAGCACAGGAGAGGA 1402
1635 TGACTTCTGTGAGCTGCCAGGCGCCGAGCATCCAGCTTAAGACCCACAGAGAGCCAGA 1694
1403 TGACCTGCTGAACCTGATCGACGCGCCCAAGTGTGTGTTGCTAAAACTGAGCAAGAGTTGA 1462
1695 GCGCAGGATGCGACGCTGTCCAGGGGAAGGGCGCTGACATCTTCCATCGGCTGGGCGC 1754
1463 CATTAATTGCCCCAGTTCTTCAAGCTGACAGTACAGTCACTTCTTCCAGAGCTGGGCGC 1522
1755 CCGTCTGTGTCTCAGCCAAAGAACCGGTGGCGGCTGGTGGGCGCGCTCCACCTGACCCG 1814
1523 CTTATCTGTGTTTTCGGCTAACAGCGTGGAGCGCTCTCTGAAAGCATCCGCTTCACTGC 1582
1815 AGGAGAGGCGGCTTTGGCCCTACGCTTCGGGAGACTCGCTCTCTCTCATCGCTGCGGT 1874
1583 AGAAGAGGGGACTTTGGGGTTTCACTTGAAGAGGAAACGCGCCCGTTCAGGTTCACTTCT 1642
1875 CATTCAGGAGAGCCAGCGCGCGCTGGGCTGAGAGGAGGCGCACTACATTTGTCTAGT 1934
1643 GGATCTTACTGCTCTGCTCGGTGGCAGGAGCCCGGAGAGAGATTATATTGCTCCAT 1702
1935 GAATGGGAGCCATGCAAGGTGTGGAGACACGCGGAGGTGTGTGAGGAGTGAAGGCTGC 1994
1703 TCAGCTTGTGATTGTAAGTGGCTGACGCTGAGTGAGGTTATGAAGCTGCTGAAGAGCTT 1762
1995 GGGAGAGCGCGCGCCAGCTGACAGTGTGTGCTGCTG 2034
1763 TGGCGAGGACGAGATCGAGATGAAGTCGTGAGCCTCTG 1802
```

Search completed: July 13, 2004, 12:02:39
Job time : 1011.06 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 13, 2004, 00:06:40 ; Search time 122.439 Seconds
(without alignments)
11190.705 Million cell updates/sec

Title: US-10-697-266-1
Perfect score: 2469
Sequence: 1 tcgcggccagtggtgcgg.....aaaaaaaaaaaaaaaaa 2469

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/prodata/2/ina/5A_COMB.seq: *
2: /cgn2_6/prodata/2/ina/5B_COMB.seq: *
3: /cgn2_6/prodata/2/ina/6A_COMB.seq: *
4: /cgn2_6/prodata/2/ina/6B_COMB.seq: *
5: /cgn2_6/prodata/2/ina/PCITUS_COMB.seq: *
6: /cgn2_6/prodata/2/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	2469	100.0	2469	4	US-09-849-334-1	Sequence 1, Appli
2	2469	100.0	2469	4	US-10-274-878-1	Sequence 1, Appli
3	946.8	38.3	1155	4	US-09-833-381-1293	Sequence 1293, Ap
4	368	14.9	18025	4	US-09-849-334-3	Sequence 3, Appli
5	368	14.9	18025	4	US-10-274-878-3	Sequence 3, Appli
6	68	2.8	2469	4	US-09-849-334-1	Sequence 1, Appli
7	68	2.8	2469	4	US-10-274-878-1	Sequence 1, Appli
8	57.6	2.3	38653	4	US-09-922-445-1	Sequence 1, Appli
9	57.2	2.3	28958	1	US-08-258-261B-6	Sequence 6, Appli
10	57.2	2.3	28958	1	US-08-456-837-6	Sequence 6, Appli
11	57.2	2.3	28958	1	US-08-457-342-6	Sequence 6, Appli
12	57.2	2.3	28958	1	US-08-457-646A-6	Sequence 6, Appli
13	57.2	2.3	28958	1	US-08-458-076A-6	Sequence 6, Appli
14	57.2	2.3	28958	1	US-08-764-233A-4	Sequence 4, Appli
15	57.2	2.3	28958	1	US-08-457-335A-6	Sequence 6, Appli
16	57.2	2.3	28958	1	US-08-729-214-6	Sequence 6, Appli
17	57.2	2.3	28958	3	US-09-028-934-6	Sequence 6, Appli
18	57.2	2.3	49377	1	US-08-764-233A-1	Sequence 1, Appli
19	54.6	2.2	7218	1	US-08-232-463-14	Sequence 14, Appli
20	52	2.1	1578	4	US-09-252-991A-5965	Sequence 5965, Ap
21	52	2.1	2469	4	US-09-252-991A-5908	Sequence 5908, Ap
22	52	2.1	3135	4	US-09-252-991A-5922	Sequence 5922, Ap
23	51.8	2.1	7065	4	US-09-874-923-115	Sequence 115, App
24	51.2	2.1	1155	4	US-09-833-381-1293	Sequence 1293, Ap
25	51	2.1	9053	4	US-09-976-594-306	Sequence 306, App
26	49.8	2.0	1992	4	US-09-252-991A-13771	Sequence 13771, A
27	49.8	2.0	2175	4	US-09-252-991A-13657	Sequence 13657, A

ALIGNMENTS

RESULT 1
US-09-849-334-1
; Sequence 1, Application US/09849334
; Patent No. 6500655
; GENERAL INFORMATION:
; APPLICANT: RUSCH, Douglas et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: THERPOF
; FILE REFERENCE: CL001099-CIP
; CURRENT APPLICATION NUMBER: US/09/849,334
; CURRENT FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2469
; TYPE: DNA
; ORGANISM: Human
US-09-849-334-1

Query Match		100.0%;	Score 2469;	DB 4;	Length 2469;
Best Local Similarity		100.0%;	Pred. No. 0;		
Matches 2469;		Conservative	0;	Mismatches	0;
				Indels	0;
				Gaps	0;
QY	1	TCGCGGCCAGTGTGTGCGGGCGGCCCTAGCCCGCTGCGAGCGCTGCGAGCGCGG	60		
DB	1	TCGCGGCCAGTGTGTGCGGGCGGCCCTAGCCCGCTGCGAGCGCTGCGAGCGCGG	60		
QY	61	GCTGCTCAGCCCGAGGAGACCCCGAGCGAGCGGCTGCGGCGATCTCTGAGGAGG	120		
DB	61	GCTGCTCAGCCCGAGGAGACCCCGAGCGAGCGGCTGCGGCGATCTCTGAGGAGG	120		
QY	121	CCGACGCGCGGGCGCGCGAGGAGCGCGCGCTGCGAGATCTAGAGAGAAACCC	180		
DB	121	CCGACGCGCGGGCGCGCGAGGAGCGCGCGCTGCGAGATCTAGAGAGAAACCC	180		
QY	181	AGGAAACACCTGTGAGCTCTTTACGGGGAAGACGGGAGGCGCTGAGAGAGTGTGCG	240		
DB	181	AGGAAACACCTGTGAGCTCTTTACGGGGAAGACGGGAGGCGCTGAGAGAGTGTGCG	240		
QY	241	TGGAGAGGCTCGGGTCCACAGAGGGGAGACCCAGTGTGTGACGTTGGCCCATG	300		
DB	241	TGGAGAGGCTCGGGTCCACAGAGGGGAGACCCAGTGTGTGACGTTGGCCCATG	300		
QY	301	AATCGGAGCTTCAATGAGTGGGTGTGACTCTCTGACGAGATCCAGTGGGCCAGCTG	360		
DB	301	AATCGGAGCTTCAATGAGTGGGTGTGACTCTCTGACGAGATCCAGTGGGCCAGCTG	360		
QY	361	CAGAGCGCGGGCCCAAGATTACACAGAGATTGACAGGAGCTGCAGATCGGACGGGC	420		
DB	361	CAGAGCGCGGGCCCAAGATTACACAGAGATTGACAGGAGCTGCAGATCGGACGGGC	420		

Sequence 13575, A
Sequence 13466, A
Sequence 11560, A
Sequence 11676, A
Sequence 50, Appl
Sequence 6162, Ap
Sequence 5434, Ap
Sequence 5508, Ap
Sequence 1, Appli
Sequence 15100, A
Sequence 14520, A
Sequence 14841, A
Sequence 22, Appl
Sequence 22, Appl
Sequence 22, Appl
Sequence 22, Appl
Sequence 1216, Ap
Sequence 14123, A

		Best Local Similarity 99.2%; Pred. No. 1.4e-180;			
		Matches 962; Conservative 0; Mismatches 7; Indels 1; Gaps 1;			
Db	1801	GTCCACCTGACCCGAGGAGGCGCTTTGGGCTTCA	CGCTTGGGAGACTCGCTGTC	1860	
QY	1861	CTCATCGCTGCGCTCATTTCCAGGAGAGCCAGGCGCGCGCTGCGCTGAGGAGGCGAC	1920		
Db	1861	CTCATGCTGCGCTCATTTCCAGGAGAGCCAGGCGCGCGCTGCGCTGAGGAGGCGAC	1920		
QY	1921	TACATTTGTGTACTGAATGGGAGCCATGAGTGTGTGAGACACGCGGAGGTGTGACG	1980		
Db	1921	TACATTTGTGTACTGAATGGGAGCCATGAGTGTGTGAGACACGCGGAGGTGTGACG	1980		
QY	1981	GAGCTGAGGCTCGGGAGAGGCGGCGCGAGCTGAGTGTGTGCTGCTGCTGCCAGC	2040		
Db	1981	GAGCTGAGGCTCGGGAGAGGCGGCGCGAGCTGAGTGTGTGCTGCTGCTGCCAGC	2040		
QY	2041	TCTAGACTGCTGCGGAGGAGCCGCGCGGCTGCTGCTGCGGCGGCTTCTA	2100		
Db	2041	TCTAGACTGCTGCGGAGGAGCCGCGCGGCTGCTGCTGCGGCGGCTTCTA	2100		
QY	2101	AGGAGCCAGAGGAGGAGTGTGAGACCCCGGCGCATCCAGTGGGCGAGTCCCGGCGC	2160		
Db	2101	AGGAGCCAGAGGAGGAGTGTGAGACCCCGGCGCATCCAGTGGGCGAGTCCCGGCGC	2160		
QY	2161	CTCTCAACTGGAGCCGAAAGGCGCCAGCAGGCGAAGACTGGAGGCTGCCCGCCAGCCCTGT	2220		
Db	2161	CTCTCAACTGGAGCCGAAAGGCGCCAGCAGGCGAAGACTGGAGGCTGCCCGCCAGCCCTGT	2220		
QY	2221	GCCCCAGTGAAGCCAGCTCCGCGCTCATCTTGAAGCACCCAGAGGCTGGCGGTGAGGCGCA	2280		
Db	2221	GCCCCAGTGAAGCCAGCTCCGCGCTCATCTTGAAGCACCCAGAGGCTGGCGGTGAGGCGCA	2280		
QY	2281	GGATCCCTGACGCGCTCAGCCCTGAGCTGAGTGGGAGCAGCAGGAGTCCCTCC	2340		
Db	2281	GGATCCCTGACGCGCTCAGCCCTGAGCTGAGTGGGAGCAGCAGGAGTCCCTCC	2340		
QY	2341	CCACCCAGAGGAGCTCCGCGCAATGCTGTCCGCGCTCATGCTGGAGGCTGCCCTCGGCGCA	2400		
Db	2341	CCACCCAGAGGAGCTCCGCGCAATGCTGTCCGCGCTCATGCTGGAGGCTGCCCTCGGCGCA	2400		
QY	2401	CCTGCTGCGCCATTAAAGACTGGTTCAGACCTGTCTGAAAAA	2460		
Db	2401	CCTGCTGCGCCATTAAAGACTGGTTCAGACCTGTCTGAAAAA	2460		
QY	2461	AAAAAAAAA 2469			
Db	2461	AAAAAAAAA 2469			

```
US-09-849-334-3
; Sequence 3, Application US/09849334
; Patent No. 6500655
; GENERAL INFORMATION:
; APPLICANT: RUSCH, Douglas et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001099-CIP
; CURRENT APPLICATION NUMBER: US/09/849,334
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 19025
; TYPE: DNA
; ORGANISM: Human
US-09-849-334-3

Query Match
Best Local Similarity 14.9%; Score 368; DB 4; Length 19025;
Matches 379; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 2058 GGGGACCGCCCGCCGCTCTGCTGGGCCCCCAGGGGGCTTCTTAAGGAGCCAGAGGGAGCA 2117
Db 15811 GGGGACCGCCCGCCGCTCTGCTGGGCCCCCAGGGGGCTTCTTAAGGAGCCAGAGGGAGCA 15870
QY 2118 TGGTTGCAAGACCCCGGCATCCACGTGGGCCAGTCCCGGCCCTCTCAACTGGAGCCG 2177
Db 15871 TGGTTGCAAGACCCCGGCATCCACGTGGGCCAGTCCCGGCCCTCTCAACTGGAGCCG 15930
QY 2178 AAAGCCCCAGGAGGCGAGACTGGAGGCTGCCCCAGCCCTGTGCCCGAGTGAAGCCAGC 2237
Db 15931 AAAGCCCCAGGAGGCGAGACTGGAGGCTGCCCCAGCCCTGTGCCCGAGTGAAGCCAGC 15990
QY 2238 TCCGCCCTCATCTTGAAGCACCCAGGGGTGGCGGTGAGGGCCAGGATCCCTGCACGCCCT 2297
Db 15991 TCCGCCCTCATCTTGAAGCACCCAGGGGTGGCGGTGAGGGCCAGGATCCCTGCACGCCCT 16049
QY 2298 CAGCCCTGGCTCCAGCTGGCAGCAAGCACCGAGCATGCCCTCCCAACCCAGAGGACCTCC 2357
Db 16050 CAGCCCTGGCTCCAGCTGGCAGCAAGCACCGAGCATGCCCTCCCAACCCAGAGGACCTCC 16109
QY 2358 GGGCAATGCTGTCCCGCTCATGCTGGAGGCTGCTGGGCGACCTGCTGCCCATTA 2417
Db 16110 GGGCAATGCTGTCCCGCTCATGCTGGAGGCTGCTGGGCGACCTGCTGCCCATTA 16169
QY 2418 GACTGGTCAGACCTGTCTGA 2437
Db 16170 GACTGGTCAGACCTGTCTGA 16189

RESULT 6
US-09-849-334-1/c
; Sequence 1, Application US/09849334
; Patent No. 6500655
; GENERAL INFORMATION:
; APPLICANT: RUSCH, Douglas et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001099-CIP
; CURRENT APPLICATION NUMBER: US/09/849,334
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2469
; TYPE: DNA
; ORGANISM: Human
US-09-849-334-1

Query Match
Best Local Similarity 2.8%; Score 68; DB 4; Length 2469;
Matches 158; Conservative 0; Mismatches 150; Indels 0; Gaps 0;

QY 1446 GCGCAGGAGCTGGAGGAGCCAGGCGAGCTTGGCAAGGCACACCTGAAGCGTGCATCCT 1505
Db 1753 GCGCAGGAGCTGGAGGAGCCAGGCGAGCTTGGCAAGGCACACCTGAAGCGTGCATCCT 1694
QY 1506 GGGGAGGAGGAGCGCTGGCGCTGACGCCCTGTGGCGGCTGCTGCCGAGGAGGAGCT 1565
Db 1693 CTGGCTTCTGGTGGGTCTTAGGCTGGATGTCGGGGCTCGGCAGCCTCACAGAGTCAT 1634
QY 1566 GCTTCGGGCTGTGATCTCCCAAGCGCTGACGCGCTCACTGGCCCAAGTATCGGAGCTCA 1625
Db 1633 CCTCAGGTCGAGCTCCGCATCTTGGCCAGTGTGAGCGCTGTGAGGATCAAG 1574
QY 1626 CCGTGAGGATGACTTCTGTGAGGCTGCCGAGGCCCGGCACATCCAGCTAAGACCCACCA 1685

US-09-849-334-3
; Sequence 3, Application US/10274878
; Patent No. 6670163
; GENERAL INFORMATION:
; APPLICANT: RUSCH, Douglas et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001099-CIP-DIV
; CURRENT APPLICATION NUMBER: US/10/274,878
; PRIOR FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: 09/849,334
; PRIOR FILING DATE: 2001-05-07
; PRIOR APPLICATION NUMBER: 09/773,371
; PRIOR FILING DATE: 2001-02-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 19025
; TYPE: DNA
```


Db 1573 CCGAAGCAGGTCACCTCGCAGACGCGGCACAGGGGTGCAGCGCGCCCTCCT 1514
QY 1686 GAAGCCAGAGCCGAGGATGACCGCTGTCCAGGGGAGGGCCCTGACATCTTCATCG 1745
Db 1513 CCGTCCCGCAGGATGCGACGCTTCAGGTGTGCTTGCCAGCTGCTCGCTCCAGCT 1454
QY 1746 GCTGGGGC 1753
Db 1453 CCGCGGC 1446

RESULT 7

US-10-274-878-1/c
; Sequence 1, Application US/10274878
; Patent No. 6670163
; GENERAL INFORMATION:
; APPLICANT: RUSCH, Douglas et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; FILE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; CURRENT APPLICATION NUMBER: US/10/274,878
; CURRENT FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: 03/849,334
; PRIOR FILING DATE: 2001-05-07/773,371
; PRIOR APPLICATION NUMBER: 09/773,371
; PRIOR FILING DATE: 2001-02-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2469
; TYPE: DNA
; ORGANISM: Human
US-10-274-878-1

Query Match 2.8%; Score 68; DB 4; Length 2469;

Best Local Similarity 51.3%; Pred. No. 8.8e-05;
Matches 158; Conservative 0; Mismatches 150; Indels 0; Gaps 0;
QY 1446 GCCGACGAGCTGAGGAGCGCAGCTTGGCAGGCACACCTGAAGCGTGCCATCCT 1505
Db 1753 GCCCAGCCGATGAGATGTACGGCCCTTCCCTGGGACAGCGGTGGCATCTCGGCT 1694
QY 1506 GGGCAGAGAGCGCTGCGCTGACGCCCTGTGCGGCTCTGCGGAGTGACCT 1565
Db 1693 CTGGCTTCTGTGGGTCTTAGGCTGGATGTCCGGGGCTCGGCAGCCTCACAGAAGTCAT 1634
QY 1566 GCTTCGGGCTGTGATCTCCACAGCGCTGCAGCGCTCACTGGCCCAAGTATGCGAGCTCGA 1625
Db 1633 CTTACGGTCAGCTCCGCTACTTGGCCAGTGAGCGCTGAGCGCTGCGGAGATCACAG 1574
QY 1626 CCGTGAGGATGACTTCTGTAGGCTGCGAGGCCCGGACATCCAGCTTAAGACCCACCA 1685
Db 1573 CCGAAGCAGCTCCACCTCGCGCAGGACGCGGCTGCGAGCGCGCAGCGCTCCT 1514
QY 1686 GAAGCCAGAGCCAGGATGACCGCTGTCCAGGGAGGGGCTGACATCTTCATCG 1745
Db 1513 CCGTCCCGCAGGATGCGACGCTTCAGGTGTGCTTTGCCAAGCTGCTCGCTCCTCCAGCT 1454
QY 1746 GCTGGGC 1753
Db 1453 CCGCGGC 1446

RESULT 8

US-09-922-445-1
; Sequence 1, Application US/09922445
; Patent No. 6528268
; GENERAL INFORMATION:
; APPLICANT: Andersson, Maria K.
; APPLICANT: Berglund, Lars G. T.
; APPLICANT: Reneland, Rikard H.

; APPLICANT: Adam, Gail I. R.
; TITLE OF INVENTION: REAGENTS AND METHODS FOR DETECTION OF HEART FAILURE
; FILE REFERENCE: G6126US
; CURRENT APPLICATION NUMBER: US/09/922,445
; CURRENT FILING DATE: 2001-08-03
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 38653
; TYPE: DNA
; ORGANISM: homo sapiens
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (1)..(26156)
; OTHER INFORMATION:
; NAME/KEY: misc feature
; LOCATION: (24801)..(24901)
; OTHER INFORMATION: nucleotide 24801 is a single nucleotide polymorphism which can be
; OTHER INFORMATION: A or G
; NAME/KEY: misc feature
; LOCATION: (24941)..(24941)
; OTHER INFORMATION: nucleotide 24941 is a single nucleotide polymorphism which can be
; OTHER INFORMATION: T or C
; NAME/KEY: exon
; LOCATION: (26157)..(26252)
; OTHER INFORMATION:
; NAME/KEY: Intron
; LOCATION: (26253)..(26401)
; OTHER INFORMATION:
; NAME/KEY: exon
; LOCATION: (26402)..(26543)
; OTHER INFORMATION:
; NAME/KEY: Intron
; LOCATION: (26544)..(27024)
; OTHER INFORMATION:
; NAME/KEY: exon
; LOCATION: (27025)..(27178)
; OTHER INFORMATION:
; NAME/KEY: Intron
; LOCATION: (27179)..(30519)
; OTHER INFORMATION:
; NAME/KEY: misc feature
; LOCATION: (27645)..(27645)
; OTHER INFORMATION: nucleotide 27645 is a single nucleotide polymorphism which can be
; OTHER INFORMATION: C or G
; NAME/KEY: exon
; LOCATION: (30520)..(30681)
; OTHER INFORMATION:
; NAME/KEY: Intron
; LOCATION: (30682)..(30894)
; OTHER INFORMATION:
; NAME/KEY: exon
; LOCATION: (30895)..(31027)
; OTHER INFORMATION:
; NAME/KEY: Intron
; LOCATION: (31028)..(31747)
; OTHER INFORMATION:
; NAME/KEY: exon
; LOCATION: (31748)..(31841)
; OTHER INFORMATION:
; NAME/KEY: Intron
; LOCATION: (31842)..(32400)
; OTHER INFORMATION:
; NAME/KEY: misc feature
; LOCATION: (32163)..(32163)
; OTHER INFORMATION: nucleotide 32163 is a single nucleotide polymorphism which can be
; OTHER INFORMATION: A or C
; NAME/KEY: exon
; LOCATION: (32401)..(32528)
; OTHER INFORMATION:
; NAME/KEY: Intron
; LOCATION: (32529)..(33414)
; OTHER INFORMATION:


```

; NAME/KEY: misc feature
; LOCATION: (32614)..(32614)
; OTHER INFORMATION: nucleotide 32614 is a single nucleotide polymorphism which can be
; NAME/KEY: exon
; LOCATION: (33415)..(33597)
; OTHER INFORMATION:
; NAME/KEY: Intron
; LOCATION: (33598)..(34314)
; OTHER INFORMATION:
; NAME/KEY: exon
; LOCATION: (34315)..(34588)
; OTHER INFORMATION:
; NAME/KEY: Intron
; LOCATION: (34589)..(36404)
; OTHER INFORMATION:
; NAME/KEY: exon
; LOCATION: (36405)..(36523)
; OTHER INFORMATION:
; NAME/KEY: Intron
; LOCATION: (36524)..(38341)
; OTHER INFORMATION:
; NAME/KEY: exon
; LOCATION: (38342)..(38653)
; OTHER INFORMATION:
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: Genbank/AC004923
; DATABASE ENTRY DATE: 1999-12-21
; RELEVANT RESIDUES: (1)..(38653)
US-09-922-445-1

```

```

Query Match          2.3%; Score 57.6; DB 4; Length 38653;
Best Local Similarity 50.7%; Pred. No. 0.02;
Matches 138; Conservative 0; Mismatches 134; Indels 0; Gaps 0;

QY 1450 CAGGAGCTGGAGAGCGCAGGAGCTTGGCAAGGCACACTGAAGCGTGCCATCTCTGGGG 1509
Db 26031 CTGGAGGAGCGCGCGCGAGCGGCGGACAGGGCGGGGGGCGCAGGTCCCGCGCCCGCG 26090

QY 1510 CAGGAGAGCGCTGCGGTGCAAGCCCTGTGCGCGTCTCTGCGGAGGTGACCTGTCTT 1569
Db 26091 GAGGAAGTGGCGGGCGGCTCGCAGGAGCTCCGGGGCGACCGCGGAGTCCGAGTAGTT 26150

QY 1570 CGGGCTGTGATCTCCAGAGCTGCGAGCGCTACTGGCCAAAGTATCGGAGCTCGACCGT 1629
Db 26151 CGGGCATAGGAGCGGAGCGCGCGCTCTTACCCGATGGCGGGGCTCGGGGGCGCAGGCG 26210

QY 1630 GAGGATGACTTGTGAGGCTGCGGAGGCGCGGACATCCAGCTAAGACCCACAGAGAG 1689
Db 26211 GACGAGGAGCTGCTCGGGGTCCCGAGCGGGCCCGGAGCCCGGTGAGCGGGGACCGGAG 26270

QY 1690 CCAGAGGCGAGATGCCAGCGCTGTCCAGGG 1721
Db 26271 TGCAGGCGCGGCGTCCCTGTGGGGCCAGGG 26302

```

```

RESULT 9
US-08-258-261B-6
; Sequence 6, Application US/08258261B
; Patent No. 5639949
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James M.
; APPLICANT: Beck, James Joseph
; APPLICANT: Hill, Dwight Steven
; APPLICANT: Ryals, John Andrew
; APPLICANT: Gaffney, Thomas Deane
; APPLICANT: Lam, Stephen Ting
; APPLICANT: Hammer, Phillip E.
; APPLICANT: Uknes, Scott Joseph
; TITLE OF INVENTION: Genes for the synthesis of
; TITLE OF INVENTION: antipathogenic substances
; NUMBER OF SEQUENCES: 22

```

```

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ciba-Geigy Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: NY
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/258,261B
; FILING DATE: 08-JUN-1994
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/457,205
; FILING DATE: 01-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Elmer, James Scott
; REGISTRATION NUMBER: 36,129
; REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8614
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28958 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-258-261B-6

```

```

Query Match          2.3%; Score 57.2; DB 1; Length 28958;
Best Local Similarity 44.8%; Pred. No. 0.023;
Matches 267; Conservative 0; Mismatches 323; Indels 6; Gaps 1;

QY 1073 CTTCCCTCTGCGCATCTGGAGCAGCTCATGATGGCCCCAGGCCAGGAATGTGTTTGGG 1132
Db 21797 CGACGATCTGCCCCCGGGAAGCGGTGCGCTGCCACCTACCCCTTCCAGCGGAGC 21856

QY 1133 GCCTCTCACCACTCGCTCCATGGCCCCCAAGACTGCTGGCCCCAGCTGGCGCTGGGCG 1192
Db 21857 GCTTCTGGCTGACCGCTTCCAGGCAACCGCGCGGGTCAACCACTTGTCTCGCTCG 21916

QY 1193 AGGAGGCGCGCCAGGTGGCGAGCTAGTACAGGCTAGTGACCGGACCATGGGCCAGCCAC 1252
Db 21917 AGGGCGGTCTTGGCAGGCGCATCGAGAGGGGAATTCGACGCGCTCAGCGGCGAGCTCC 21976

QY 1253 CGTCCACGACTACGTGCTGTCTCTGGAATGCGCTGGTGCATGTCAAGGCCGAGTACT 1312
Db 21977 ACGTGACGGGACGAGCAGCGCGCGCTTGCCTTCTCTCCACCTCTGGAGCT 22036

QY 1313 TCCGCTCCCTGGCCCACTACCACTAGCATGGCCCTCTGCGAGCGCTCCCGAGGACCG 1372
Db 22037 TTCGCCAGCGGGCAAGAGCAGGCGACGGTTCGAGCGCTGGGCTTACCGCATCACGTGA 22096

QY 1373 AGGAGAGCTCTCCCAACGACAGCAGAGGTCTTCTTGCAGCCCCCCCCACCTCTTAAGCCCC 1432
Db 22097 AGCTCTGACCAACCGCCACACGCGCCCGGCGGAGCTTGG-----CGGCACTTGGCTCTCG 22150

QY 1433 GAGGCCCTGTGCTCGCGCAGAGCTGGAGGAGCGCAGCAGCTTGGCAAGGCGCACCTGA 1492
Db 22151 TCGTGGCGGCGCTCTGGACGACGCGCTCCCTCCGCGCTCACCGAGCGGCTCGGCC 22210

QY 1493 AGCGTGCCATCTCGGGGACGAGGAGGCGCTGCGGCTGACGCGCTGTGCGCGCTCTGCG 1552
Db 22211 GCGCGGCGCGCGCTCTCTCGCGCTGAGCCAGGGCCACCTGGACCGGAGGCTC 22270

```


Db	22097	AGCCTGTGACCAACCGCCACACGCGCGGAGCTGG-----CCGGCACCTGGCTCTCTCG	22150
QY	1433	GAGGCCCTGTGTCGCGCAGGAGCTGGAGGACGCGAGGCGAGCTTGGCAAGGCACACCTGA	1492
Db	22151	TGCTGCCGCGCGCTCTGGACGACGACGCGCTCCCTCCGCGCTCACCGAGGCGCTCGGCC	22210
QY	1493	AGCGTGCATCTCTGGGGCAGGAGGCGCTCGGCTGCACGCCCTGTGCGCGTCTCTGC	1552
Db	22211	GGCGGGCGGCGCTCTCGCGTGGCTGAGCCACCTGACCGGAGGCTC	22270
QY	1553	GCAGGTGACCTGCTCTTCGGGCTGTGATCTCCAGCGCTGCAGCGCTCATCTGGCCAAGT	1612
Db	22271	TCCCGAGCACTGCGCCAGGCTTGCGCGGAGACCGCGCGCTCGCGCGTGTCTCGC	22330
QY	1613	ATCGGAGGCTCGACCGTGAGGATGATCTTGTGAGGCTGCCGAGGCCCGGACATC	1668
Db	22331	TCCTGCTCTTCGACGAAGTCCCTCGCGACCATGTGCGCCCTGCGCCGCGGAGCTC	22386

```

RESULT 13
US-08-458-076A-6
; Sequence 6, Application US/08458076A
; Patent No. 5698425
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas M.
; APPLICANT: Ligon, James M.
; APPLICANT: Beck, James Joseph
; APPLICANT: Hill, Dwight Steven
; APPLICANT: Ryals, John Andrew
; APPLICANT: Gaffney, Thomas Deane
; APPLICANT: Lam, Stephen Ting
; APPLICANT: Hammer, Phillip E.
; APPLICANT: Uknes, Scott Joseph
; TITLE OF INVENTION: Genes for the synthesis of
; TITLE OF INVENTION: antipathogenic substances
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ciba-Geigy Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: NY
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/458,076A
; FILING DATE: 01-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/457,205
; FILING DATE: 01-JUN-1995
; APPLICATION NUMBER: 08/258,261
; FILING DATE: 08-Jun-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Elmer, James Scott
; REGISTRATION NUMBER: 36,129
; REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8614
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28958 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO

```

;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/764,233A
;; FILING DATE:
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/08/729,214
;; FILING DATE: 09-OCT-1996
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/08/258,261
;; FILING DATE: 08-JUN-1994
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Meigs, J. Timothy
;; REGISTRATION NUMBER: 38,241
;; REFERENCE/DOCKET NUMBER: 1506/CIP6
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (919) 541-8587
;; TELEFAX: (919) 541-8689
;; INFORMATION FOR SEQ ID NO: 4:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 28958 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
;; ORIGINAL SOURCE:
;; ORGANISM: Sorangium cellulosum
;; IMMEDIATE SOURCE:
;; CLONE: p98/1
;; US-08-764-233A-4

Query Match 2.3%; Score 57.2; DB 1; Length 28958;
Best Local Similarity 44.8%; Pred. No. 0.023;
Matches 267; Conservative 0; Mismatches 323; Indels 6; Gaps 1;
QY 1073 CGTCTCTTCGCACTGAGCAGCTCATGATGCCCGCCAGGACCCAGGAATGTGTTGAGG 1132
Db |||||
QY 21797 CGACGATCTGCCCCCGGAGCGGTGCCCTGCCACCTACCCCTTCAGCGGAGC 21856
Db |||||
QY 1133 GCCTCTCACCCTGCTCCATGATGCCCGCCAGGACTGCTGGCCCGCAGCTGGCCGCGC 1192
Db |||||
QY 21857 GCTTCTGGCTCGACGCTCCACGCGCCCGCGCGCTCAACCACTTGTCTCGCTCG 21916
Db |||||
QY 1193 AGGAGCGCGCCAGGTGCGACGCCAGTACAGGCTAGTGACCGGACCATGGCCCGACAC 1252
Db |||||
QY 21917 AGGGCGGTTCTGGCAGGCGCATCGAGCGGGGAATATCGACCGCTCAGCGCGCAGCTCC 21976
Db |||||
QY 1253 CGTCCACGACTACGTGCTCTCTGACTGCGCTTGTGTCATGTCAAGGCCGAGTACT 1312
Db |||||
QY 21977 AGTGGACGGGACGAGCAGCGCGCGCTTGCCTGCTCTTCCACCTTCGCGAGCT 22036
Db |||||
QY 1313 TCCGCTCCGTGGCCCACTACACGTAGCCATGGCCCTCTGAGCGGCTCCCGAGCGACCG 1372
Db |||||
QY 22037 TTCGCGCAGCGCGCAAGCAGCGGCGCTGAGCGGCTGCGGCTACCGCATCATCGTGA 22096
Db |||||
QY 1373 AGGAGAGCTCCCGACGACGAGAGGTTCTTCTGAGCGCCCGCCAGCTTGGCAAGGCACACCTGA 1492
Db |||||
QY 22097 AGCTCTGACACACCGCACCGCGCGCGCGCTG-----CCGGCACCTGGCTCTCG 22150
Db |||||
QY 1433 GAGGCCCTGTGTGTCGCGCAGGAGCTGGAGGCGCGAGGCTGGCAAGGCACACCTGA 1492
Db |||||
QY 22151 TCGTCCGCGCGCTCTGGACGACGAGCGCTCCCTTCGCGCTTCACGAGGCGCTCGGCC 22210
Db |||||
QY 1493 AGCGTGCCATCTTGGGCGAGGAGGCGCTGCGCTGACGCGCTGTGCGCGCTCTGCG 1552
Db |||||
QY 22211 GCGCGCGCGCGCGCTCTCTGCGCTGCGCTGAGCGAGCGCCACCTGGACCGCGAGGCTC 22270
Db |||||
QY 1553 GCGAGGTGACCTGCTCGGGCTGTGATCTCCACACGCTGCGAGGCTCATCTGGCCAACT 1612
Db |||||
QY 22271 TCGCGAGACCTGGCGCAGGCTTGGCGCGAGACCGCGCGCGCTGCGCTGCTCTGCG 22330
Db |||||
QY 1613 ATCGGAGCTCGACCGTGAAGATGACTTCTGTGAGGCTGCGAGCGCCCGGACATC 1668
Db |||||
QY 22331 TCCTCGCCTCGACGAAGTCCCCTCGCGGACCATGCGCGCTGCGCGCGGAGCTC 22386
Db |||||

RESULT 15
US-08-457-335A-6
;; Sequence 6, Application US/08457335A
;; Patent No. 5723759
;; GENERAL INFORMATION:
;; APPLICANT: Schupp, Thomas
;; APPLICANT: Ligon, James M.
;; APPLICANT: Beck, James Joseph
;; APPLICANT: Hill, Dwight Steven
;; APPLICANT: Ryals, John Andrew
;; APPLICANT: Gaffney, Thomas Deane
;; APPLICANT: Lam, Stephen Ting
;; APPLICANT: Hammer, Phillip E.
;; APPLICANT: Uknes, Scott Joseph
;; TITLE OF INVENTION: Genes for the synthesis of
;; TITLE OF INVENTION: antipathogenic substances
;; NUMBER OF SEQUENCES: 22
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Ciba-Geigy Corporation
;; STREET: 7 Skyline Drive
;; CITY: Hawthorne
;; STATE: NY
;; COUNTRY: USA
;; ZIP: 10532
;; COMPUTER READABLE FORM:
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent In Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/457,335A
;; FILING DATE: 01-JUN-1995
;; CLASSIFICATION: 800
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/08/457,205
;; FILING DATE: 01-JUN-1995
;; APPLICATION NUMBER: 08/258,261
;; FILING DATE: 08-Jun-1994
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Elmer, James Scott
;; REGISTRATION NUMBER: 36,129
;; REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 919-541-8614
;; TELEFAX: 919-541-8689
;; INFORMATION FOR SEQ ID NO: 6:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 28958 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
;; HYPOTHETICAL: NO
;; ANTI-SENSE: NO
;; US-08-457-335A-6

Query Match 2.3%; Score 57.2; DB 1; Length 28958;
Best Local Similarity 44.8%; Pred. No. 0.023;
Matches 267; Conservative 0; Mismatches 323; Indels 6; Gaps 1;
QY 1073 CGTCTCTTCGCACTGAGCAGCTCATGATGCCCGCCAGGACCCAGGAATGTGTTGAGG 1132
Db |||||
QY 21797 CGACGATCTGCCCCCGGAGCGGTGCCCTGCCACCTACCCCTTCAGCGGAGC 21856
Db |||||
QY 1133 GCCTCTCACCCTGCTCCATGATGCCCGCCAGGACTGCTGGCCCGCAGCTGGCCGCGC 1192
Db |||||
QY 21857 GCTTCTGGCTCGACGCTCCACGCGCCCGCGCGCTCAACCACTTGTCTCGCTCG 21916
Db |||||
QY 1193 AGGAGCGCGCCAGGTGCGACGCCAGTACAGGCTAGTGACCGGACCATGGCCCGACAC 1252
Db |||||
QY 21917 AGGGCGGTTCTGGCAGGCGCATCGAGCGGGGAATATCGACCGCTCAGCGCGCAGCTCC 21976
Db |||||

Tue Jul 13 15:19:49 2004

QY	1253	CGGTCCAAGCACTACGTGCTCTCTCTGGACGTGCCCTGGTGCATGTCAAGCCCGAGTACT	1312
DB	21977	ACGTGGACGGCACAGCAGCAGCGCGCGCTTGCCCTGCTCTCTCCACCCCTCCGAGCT	22036
QY	1313	TCCGCTCCCTGGCCCACTACCACTACAGCTAGCATGGCCCTCTGCGAGGGCTCCCCACGACCG	1372
DB	22037	TTCCGCACAGCGCGCAGAGCAGGGCACGGTCTGACGCTGGCGGTACCGCATCACTGGGA	22096
QY	1373	AGGGAGAGTTCGCCACGCAACGAGCAGGTCTTCTTCAGCCCCCACTCTCTTAAGCCCC	1432
DB	22097	AGCCTCTGACCAACCGCCACCAACGCGCGGACCTGG-----CGGGCACTTGGTCTCTCG	22150
QY	1433	GAGGCCCTGTGTGCGCGCAGGAGCTGTGAGAGCGCAGGAGGTTGGCAAGCGCACACTGA	1492
DB	22151	TCTGTGCGGCGCGCTCTGTGACGACAGCGCTCTCCCTCCGCGCTCACAGAGCGTCTCGCC	22210
QY	1493	AGCGTGCCATCTTGGGGCAGGAGGCGCTCGCGCTGCAGCCCTGTGCGCGTCTCTGC	1552
DB	22211	GGCGCGCGCGCGCTCTTCGCGTGGCGCTTCTGAGCCAGGCCCACTGAGACCGCAGGCTC	22270
QY	1553	GCAGGTGGACCTGCTTTCGGGGTGTGATCTCCACAGACGCTGCAGCGCTCACTGGCCCAAGT	1612
DB	22271	TGCCCGAGCACCTGCCCGAGGCTTGGCGCGAGACCGCGCGCTTCGCGCGTGTCTCGC	22330
QY	1613	ATCCGAGCTCGAACCTGTAGGATGATCTTGTGAGGCTGCCAGGCCCCCGACATC	1668
DB	22331	TCCTCGCCCTTCGACGAAGTTCCTCCGCGACCATCGCCCGTGCCTGCCCGGAGTCT	22386

Search completed: July 13, 2004, 08:26:32
Job time : 129.439 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 13, 2004, 12:07:02 ; Search time 427 Seconds
(without alignments)
528.143 Million cell updates/sec

Title: US-10-697-266-2
Perfect score: 3774
Sequence: 1 MILERPPDAGAGEESPRLO.....QPCAPVKPAPPSLKHPGW 723

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1279676 seqs, 311918243 residues

Total number of hits satisfying chosen parameters: 1279676

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3774	100.0	723	12	US-10-697-266-2
2	3774	100.0	723	14	US-10-274-878-2
3	3449.5	91.4	670	12	US-10-451-207-2
4	3435	91.0	669	12	US-10-112-944-318
5	2331	61.8	634	12	US-10-697-266-4
6	2331	61.8	634	14	US-10-274-878-4
7	2331	61.8	643	15	US-10-239-607-71
8	1442.5	38.2	311	16	US-10-408-765A-2465
9	1278	33.9	257	9	US-09-925-301-1016
10	1278	33.9	257	14	US-10-023-896-59
11	1278	33.9	257	14	US-10-023-896-87
12	1256	33.3	686	15	US-10-108-260A-4840
13	1255	33.3	611	12	US-10-092-900A-352
14	1255	33.3	686	15	US-10-239-607-53
15	1249	33.1	686	9	US-09-895-040A-3

16	1242	32.9	686	15	US-10-239-607-73	Sequence 73, Appl
17	1228	32.5	686	16	US-10-408-765A-2644	Sequence 2644, Ap
18	1195	31.7	686	12	US-10-092-900A-350	Sequence 350, App
19	1023	27.1	718	15	US-10-239-607-72	Sequence 72, Appl
20	936	24.8	180	12	US-10-112-944-776	Sequence 776, App
21	849	22.5	351	15	US-10-378-029-101	Sequence 101, App
22	394	10.4	298	15	US-10-108-260A-2912	Sequence 2912, Ap
23	387	10.3	1403	12	US-10-425-114-54518	Sequence 54518, A
24	380.5	10.1	939	15	US-10-094-749-1895	Sequence 1895, Ap
25	380.5	10.1	1636	16	US-10-433-794-1	Sequence 1, Appl
26	301	8.0	877	12	US-10-425-114-59396	Sequence 59396, A
27	294.5	7.8	883	16	US-10-437-963-180574	Sequence 180574, A
28	274.5	7.3	868	12	US-10-221-278-251	Sequence 251, App
29	274.5	7.3	868	15	US-10-291-172-251	Sequence 251, App
30	268	7.1	921	9	US-09-800-729-199	Sequence 199, App
31	268	7.1	921	15	US-10-264-049-3015	Sequence 3015, Ap
32	264.5	7.0	868	9	US-09-800-729-106	Sequence 106, App
33	264.5	7.0	868	11	US-09-833-245-2179	Sequence 2179, Ap
34	251.5	6.7	200	12	US-10-424-599-275853	Sequence 275853, A
35	251.5	6.7	397	16	US-10-437-963-123181	Sequence 123181, A
36	232	6.1	661	9	US-09-801-368-166	Sequence 166, App
37	185	4.9	104	15	US-10-630-590-187	Sequence 187, App
38	156.5	4.1	780	9	US-09-771-161A-144	Sequence 144, App
39	156.5	4.1	942	14	US-10-354-358-74	Sequence 74, Appl
40	155.5	4.1	942	9	US-09-771-161A-235	Sequence 235, App
41	155.5	4.1	942	9	US-09-771-161A-236	Sequence 236, App
42	146	3.9	1151	15	US-10-369-493-7865	Sequence 7865, Ap
43	145.5	3.9	1522	14	US-10-134-102-8	Sequence 8, Appl
44	145.5	3.9	1522	15	US-10-144-198-31	Sequence 31, Appl
45	144.5	3.8	77	12	US-10-042-865-86	Sequence 86, Appl

ALIGNMENTS

RESULT 1
US-10-697-266-2
; Sequence 2, Application US/10697266
; Publication No. US20040067522A1
; GENERAL INFORMATION:
; APPLICANT: RUSCH, Douglas et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001039-CIP-DIV2
; CURRENT APPLICATION NUMBER: US/10/697,266
; CURRENT FILING DATE: 2003-10-31
; PRIOR FILING DATE: 2003-10-31
; PRIOR FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: 09/849,334
; PRIOR FILING DATE: 2001-05-07
; PRIOR APPLICATION NUMBER: 09/773,371
; PRIOR FILING DATE: 2001-02-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 723
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-697-266-2

Query Match	100.0%	Score 3774;	DB 12;	Length 723;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 723;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
QY	1	MILERPPDAGAGEESPRLOISRRPRKTRVSSLRGRREGLRDVCAMRGCRVHRGEDPVR	60	
Db	1	MILERPPDAGAGEESPRLOISRRPRKTRVSSLRGRREGLRDVCAMRGCRVHRGEDPVR	60	
QY	61	VHVGPMPQLHAVGCDLSLTQCGQLQSRRAQIHQIQKELQMRGTGAENLYRATSNRRVR	120	
Db	61	VHVGPMPQLHAVGCDLSLTQCGQLQSRRAQIHQIQKELQMRGTGAENLYRATSNRRVR	120	

```
QY 121 ETVALELSYNSNQLLKEELEELSGGVDPRHSGSEAVTVPMIPLGLKETKELDWSTPLK 180
DB 121 ETVALELSYNSNQLLKEELEELSGGVDPRHSGSEAVTVPMIPLGLKETKELDWSTPLK 180
QY 181 ELISVHFEGDGASYEAEIRELEALRQAMRTPSRNESGLELLTAYYNQCFDARFLTPAR 240
DB 181 ELISVHFEGDGASYEAEIRELEALRQAMRTPSRNESGLELLTAYYNQCFDARFLTPAR 240
QY 241 SLGLFFHWYDSLTGVPAAQORALAFKESVLFNIGALHTQIGARODRSCTEGARAMEAFQ 300
DB 241 SLGLFFHWYDSLTGVPAAQORALAFKESVLFNIGALHTQIGARODRSCTEGARAMEAFQ 300
QY 301 RAAGAFSLLENFNSHAPSPDMSAASLCALEQLMMAQAQECVFEGLSPASMAPQDCLAQL 360
DB 301 RAAGAFSLLENFNSHAPSPDMSAASLCALEQLMMAQAQECVFEGLSPASMAPQDCLAQL 360
QY 361 RLAQEAQAQAAEYRLVHRTMAQPPVHDYVPVSWTALVHVKAIEYFRSLAHYHVAMALCDGS 420
DB 361 RLAQEAQAQAAEYRLVHRTMAQPPVHDYVPVSWTALVHVKAIEYFRSLAHYHVAMALCDGS 420
QY 421 PATEGELTHQVFLQPTTSKPRGPVLPOELEERRQLGKAHLKRAILGQBEALRLHALC 480
DB 421 PATEGELTHQVFLQPTTSKPRGPVLPOELEERRQLGKAHLKRAILGQBEALRLHALC 480
QY 481 RVLREVDLLRAVISQTLQSLAKYAEALDREDDFCEAAEAPDIQPKTHOKPEARMPLRSQ 540
DB 481 RVLREVDLLRAVISQTLQSLAKYAEALDREDDFCEAAEAPDIQPKTHOKPEARMPLRSQ 540
QY 541 KGPDI FHLRGLSVPSAKNRWLVGPVHLTRGEGGFLTRGDSVPLTAAVIPGSOAAAA 600
DB 541 KGPDI FHLRGLSVPSAKNRWLVGPVHLTRGEGGFLTRGDSVPLTAAVIPGSOAAAA 600
QY 601 GLKEGDIYVSVNGQPCRWWRHAEVVTTELKAAEAGASLQVWSLLPSSRLPSLGDRRPVLL 660
DB 601 GLKEGDIYVSVNGQPCRWWRHAEVVTTELKAAEAGASLQVWSLLPSSRLPSLGDRRPVLL 660
QY 661 GPRGLLRSQREHGCKTPTASTWASPRLLNWSRKAQGGTGGCPQPCAPVKPAPPSLKH 720
DB 661 GPRGLLRSQREHGCKTPTASTWASPRLLNWSRKAQGGTGGCPQPCAPVKPAPPSLKH 720
QY 721 GWP 723
DB 721 GWP 723
```

RESULT 2

```
US-10-274-878-2
; Sequence 2, Application US/10274878
; Publication NO. US20030049792A1
; GENERAL INFORMATION:
; APPLICANT: RUSCH, Douglas et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001099-CIP-DIV
; CURRENT APPLICATION NUMBER: US/10/274,878
; CURRENT FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: 09/849,334
; PRIOR FILING DATE: 2001-05-07
; PRIOR APPLICATION NUMBER: 09/773,371
; PRIOR FILING DATE: 2001-02-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 2
; LENGTH: 723
; TYPE: PRT
; ORGANISM: Human
US-10-274-878-2
```

Query Match 100.0%; Score 3774; DB 14; Length 723;
Best Local Similarity 100.0%; Pred No. 0;
Matches 723; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MILERPDAAGAGBESRRLQISRRPKRTRVSSLRGRREGRLDYCAMRGCKVRHGDDPVR 60
DB 1 MILERPDAAGAGBESRRLQISRRPKRTRVSSLRGRREGRLDYCAMRGCKVRHGDDPVR 60
QY 61 VHVCPMPOLHAVGCDLSLTQCCGLOSRRAQIHQQIDKELQMRITGAENLYRATSNRVR 120
DB 61 VHVCPMPOLHAVGCDLSLTQCCGLOSRRAQIHQQIDKELQMRITGAENLYRATSNRVR 120
QY 121 ETVALELSYNSNQLLKEELEELSGGVDPRHSGSEAVTVPMIPLGLKETKELDWSTPLK 180
DB 121 ETVALELSYNSNQLLKEELEELSGGVDPRHSGSEAVTVPMIPLGLKETKELDWSTPLK 180
QY 181 ELISVHFEGDGASYEAEIRELEALRQAMRTPSRNESGLELLTAYYNQCFDARFLTPAR 240
DB 181 ELISVHFEGDGASYEAEIRELEALRQAMRTPSRNESGLELLTAYYNQCFDARFLTPAR 240
QY 241 SLGLFFHWYDSLTGVPAAQORALAFKESVLFNIGALHTQIGARODRSCTEGARAMEAFQ 300
DB 241 SLGLFFHWYDSLTGVPAAQORALAFKESVLFNIGALHTQIGARODRSCTEGARAMEAFQ 300
QY 301 RAAGAFSLLENFNSHAPSPDMSAASLCALEQLMMAQAQECVFEGLSPASMAPQDCLAQL 360
DB 301 RAAGAFSLLENFNSHAPSPDMSAASLCALEQLMMAQAQECVFEGLSPASMAPQDCLAQL 360
QY 361 RLAQEAQAQAAEYRLVHRTMAQPPVHDYVPVSWTALVHVKAIEYFRSLAHYHVAMALCDGS 420
DB 361 RLAQEAQAQAAEYRLVHRTMAQPPVHDYVPVSWTALVHVKAIEYFRSLAHYHVAMALCDGS 420
QY 421 PATEGELTHQVFLQPTTSKPRGPVLPOELEERRQLGKAHLKRAILGQBEALRLHALC 480
DB 421 PATEGELTHQVFLQPTTSKPRGPVLPOELEERRQLGKAHLKRAILGQBEALRLHALC 480
QY 481 RVLREVDLLRAVISQTLQSLAKYAEALDREDDFCEAAEAPDIQPKTHOKPEARMPLRSQ 540
DB 481 RVLREVDLLRAVISQTLQSLAKYAEALDREDDFCEAAEAPDIQPKTHOKPEARMPLRSQ 540
QY 541 KGPDI FHLRGLSVPSAKNRWLVGPVHLTRGEGGFLTRGDSVPLTAAVIPGSOAAAA 600
DB 541 KGPDI FHLRGLSVPSAKNRWLVGPVHLTRGEGGFLTRGDSVPLTAAVIPGSOAAAA 600
QY 601 GLKEGDIYVSVNGQPCRWWRHAEVVTTELKAAEAGASLQVWSLLPSSRLPSLGDRRPVLL 660
DB 601 GLKEGDIYVSVNGQPCRWWRHAEVVTTELKAAEAGASLQVWSLLPSSRLPSLGDRRPVLL 660
QY 661 GPRGLLRSQREHGCKTPTASTWASPRLLNWSRKAQGGTGGCPQPCAPVKPAPPSLKH 720
DB 661 GPRGLLRSQREHGCKTPTASTWASPRLLNWSRKAQGGTGGCPQPCAPVKPAPPSLKH 720
QY 721 GWP 723
DB 721 GWP 723
```

RESULT 3

```
US-10-451-207-2
; Sequence 2, Application US/104511207
; Publication NO. US20040038267A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: BURFORD, Neil
; APPLICANT: DING, Li
; APPLICANT: YUE, Henry
; APPLICANT: THORNTON, Michael B.
; APPLICANT: CHAWLA, Narinder K.
; APPLICANT: GANDHI, Ameena R.
; APPLICANT: ARVIZU, Chandra S.
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: SWARNAKAR, Anita
; APPLICANT: DUGGAN, Brendan M.
; APPLICANT: LU, Dyung Aina M.
; APPLICANT: THANGAVELU, Kavitha
; APPLICANT: WARREN, Bridget A.
; APPLICANT: TANG, Y. Tom
```


APPLICANT: KHAN, Farrah A.
APPLICANT: YAO, Monique G.
APPLICANT: EVERLING, Brooke M.
TITLE OF INVENTION: INTRACELLULAR SIGNALING MOLECULES
FILE REFERENCE: PF-0868 USN
CURRENT FILING DATE: 2003-06-18
PRIOR APPLICATION NUMBER: PCT/US01/50315
PRIOR FILING DATE: 2001-12-19
PRIOR APPLICATION NUMBER: US 60/257,804
PRIOR FILING DATE: 2000-12-21
PRIOR APPLICATION NUMBER: US 60/260,102
PRIOR FILING DATE: 2001-01-05
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PERL Program
SEQ ID NO 2
LENGTH: 670
TYPE: PRT
ORGANISM: Homo sapiens
NAME/KEY: misc_feature
OTHER INFORMATION: Incyte ID No. US20040038267A1 5560161CD1
US-10-451-207-2

Query Match 91.4%; Score 3449.5; DB 12; Length 670;
Best Local Similarity 92.7%; Pred. No. 1.9e-293;
Matches 670; Conservative 0; Mismatches 0; Indels 53; Gaps 1;

QY 1 MILERRPGAGAGERSPRIQISRRKPRKTRVSSLRGRREGRLDVCAMRGCRVHRGDDPVR 60
DB 1 MILERRPGAGAGERSPRIQ----- 20
QY 61 VHVGPMPQLHAGVCDLSLTQICGQLOSRAQIHQOIDKELQMTGAENLYRATSNRVR 120
DB 21 -----GCDLSLTQICGQLOSRAQIHQOIDKELQMTGAENLYRATSNRVR 67
QY 121 ETVALELSYVNSNLOLLKEELELSGGVDPGRHSGEAVTVPMIPLGKTKELDWSTPLK 180
DB 68 ETVALELSYVNSNLOLLKEELELSGGVDPGRHSGEAVTVPMIPLGKTKELDWSTPLK 127
QY 181 ELISVHFGEDGASYEAEIRELEALRQAMRTPSRNSGLELLTAYYNQLCFLDAREFLTAR 240
DB 128 ELISVHFGEDGASYEAEIRELEALRQAMRTPSRNSGLELLTAYYNQLCFLDAREFLTAR 187
QY 241 SGLGPHFWYDSLTGTPAQORALAFEGSVLFNIGALHTQIGARQDRSCTEGARRAMEAFQ 300
DB 189 SGLGPHFWYDSLTGTPAQORALAFEGSVLFNIGALHTQIGARQDRSCTEGARRAMEAFQ 247
QY 301 RAAGAFSLRENFSHAPSPDMSAASLCALQELMMAQAQECVPEGLSPAPMAPODCLAQL 360
DB 248 RAAGAFSLRENFSHAPSPDMSAASLCALQELMMAQAQECVPEGLSPAPMAPODCLAQL 307
QY 361 RLQAQAAQVAAYRVLVHRTMAQPPVHDYVPVSWTALVHVKAIFYRSLAHYHVAMALCDGS 420
DB 308 RLQAQAAQVAAYRVLVHRTMAQPPVHDYVPVSWTALVHVKAIFYRSLAHYHVAMALCDGS 367
QY 421 PATEGELPHEOVFLOPPTSSKRGVLPQELERROQLKAHLKRALIQOEALRIHALC 480
DB 368 PATEGELPHEOVFLOPPTSSKRGVLPQELERROQLKAHLKRALIQOEALRIHALC 427
QY 481 RVLEVDLLRAVISQTLQORSLAKYAELDRDDFCEAAEAPDIQPKTHQKPEARMFRLSOG 540
DB 428 RVLEVDLLRAVISQTLQORSLAKYAELDRDDFCEAAEAPDIQPKTHQKPEARMFRLSOG 487
QY 541 KGPDI FHLRGPLSVFSAXNRWLVGPVHLTRGEGGFLTRGDSVPVLI AAVIPGSOAAAA 600
DB 488 KGPDI FHLRGPLSVFSAXNRWLVGPVHLTRGEGGFLTRGDSVPVLI AAVIPGSOAAAA 547
QY 601 GLKEGDIYVSVNGQPCRRWRHAEVVTTELKAAEGASLQVSVLLPSSRLPSIGDRRPVLL 660
DB 548 GLKEGDIYVSVNGQPCRRWRHAEVVTTELKAAEGASLQVSVLLPSSRLPSIGDRRPVLL 607
QY 661 GPRGLLRSOREHGCKTPASTWASPRLLNWSRKAQOGKGTGGCPQPCAPVKPAPPSSLKHP 720

DB 608 GPRGLLRSOREHGCKTPASTWASPRLLNWSRKAQOGKGTGGCPQPCAPVKPAPPSSLKHP 667
QY 721 GWP 723
DB 668 GWP 670
RESULT 4
US-10-112-944-318
Sequence 318, Application US/10112944
Publication No. US20040048249A1
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Yang, Yonghong
APPLICANT: Weng, Gezhi
APPLICANT: Zhang, Jie
APPLICANT: Ren, Feiyan
APPLICANT: Xue, Aidong J.
APPLICANT: Wang, Jian-Rui
APPLICANT: Wehrman, Tom
APPLICANT: Ghosh, Malabika
APPLICANT: Wang, Dunrui
APPLICANT: Zhao, Qing A.
APPLICANT: Wang, Zhiwei
TITLE OF INVENTION: No. US20040048249A1el Nucleic Acids and
TITLE OF INVENTION: Secreted Polypeptides
FILE REFERENCE: 805A
CURRENT APPLICATION NUMBER: US/10/112,944
PRIOR FILING DATE: 2002-03-28
PRIOR APPLICATION NUMBER: US 09/488,725
PRIOR FILING DATE: 2000-01-21
PRIOR APPLICATION NUMBER: US 09/491,404
PRIOR FILING DATE: 2000-01-25
PRIOR APPLICATION NUMBER: US 09/496,914
PRIOR FILING DATE: 2000-02-03
PRIOR APPLICATION NUMBER: US 09/515,126
PRIOR FILING DATE: 2000-02-28
PRIOR APPLICATION NUMBER: US 09/519,705
PRIOR FILING DATE: 2000-03-07
PRIOR APPLICATION NUMBER: US 09/540,217
PRIOR FILING DATE: 2000-03-31
PRIOR APPLICATION NUMBER: US 09/552,929
PRIOR FILING DATE: 2000-04-18
PRIOR APPLICATION NUMBER: US 09/577,408
PRIOR FILING DATE: 2000-05-18
NUMBER OF SEQ ID NOS: 924
SOFTWARE: pt_FL_genes Version 5.0
SEQ ID NO 318
LENGTH: 669
TYPE: PRT
ORGANISM: Homo sapiens
US-10-112-944-318

Query Match 91.0%; Score 3435; DB 12; Length 669;
Best Local Similarity 92.5%; Pred. No. 3.6e-292;
Matches 669; Conservative 0; Mismatches 0; Indels 54; Gaps 2;

QY 1 MILERRPGAGAGERSPRIQISRRKPRKTRVSSLRGRREGRLDVCAMRGCRVHRGDDPVR 60
DB 1 MILERRPGAGAGERSPRIQ----- 20
QY 61 VHVGPMPQLHAGVCDLSLTQICGQLOSRAQIHQOIDKELQMTGAENLYRATSNRVR 120
DB 21 -----GCDLSLTQICGQLOSRAQIHQOIDKELQMTGAENLYRATSNRVR 67
QY 121 ETVALELSYVNSNLOLLKEELELSGGVDPGRHSGEAVTVPMIPLGKTKELDWSTPLK 180
DB 68 ETVALELSYVNSNLOLLKEELELSGGVDPGRHSGEAVTVPMIPLGKTKELDWSTPLK 127
QY 181 ELISVHFGEDGASYEAEIRELEALRQAMRTPSRNSGLELLTAYYNQLCFLDAREFLTAR 240
DB 128 ELISVHFGEDGASYEAEIRELEALRQAMRTPSRNSGLELLTAYYNQLCFLDAREFLTAR 187

```
QY 241 SLGLFFHWYDSLTCGPAQORALAFKESVLFNIGALHTQIGARODRSCTEGARRAMEAFQ 300
DB 188 SLGLFFHWYDSLTCGPAQORALAFKESVLFNIGALHTQIGARODRSCTEGARRAMEAFQ 247
QY 301 RAAGAFSLIRENFHSHAPSPDMSAASLCALEQLMMAQAQECVFEGLSPPASMAPQDCIAQL 360
DB 248 RAAGAFSLIRENFHSHAPSPDMSAASLCALEQLMMAQAQECVFEGLSPPASMAPQDCIAQL 307
QY 361 RLAGEAAQVAAEYRLVHRTMAQPPVHDYVPSWTALVHVKAAYFRSLAHYHVAMALCDGS 420
DB 308 RLAGEAAQVAAEYRLVHRTMAQPPVHDYVPSWTALVHVKAAYFRSLAHYHVAMALCDGS 367
QY 421 PATEGELPTHQVLOPFTSSKPRGPVLPQLEERRQLGKAHLKRAILGQBEALRLHALC 480
DB 368 P-TATEGELPTHQVLOPFTSSKPRGPVLPQLEERRQLGKAHLKRAILGQBEALRLHALC 426
QY 481 RVLREVDLLRAVISQTLQRSIAKYAELDREDDFCEAAEAPDIQPKTHQKPEARMPLRSQ 540
DB 427 RVLREVDLLRAVISQTLQRSIAKYAELDREDDFCEAAEAPDIQPKTHQKPEARMPLRSQ 486
QY 541 KGPDIHRLGLPLSVFSAKNRWLVGPVHLTRGEGGFGTLRGDSPLVIAAVIPGSAQAAA 600
DB 487 KGPDIHRLGLPLSVFSAKNRWLVGPVHLTRGEGGFGTLRGDSPLVIAAVIPGSAQAAA 546
QY 601 GLKEGDYIVSVNGQPCRWWRHAEVVTTELKAAEAGASIQVWSLLPSSRLPSLGDRRPVLL 660
DB 547 GLKEGDYIVSVNGQPCRWWRHAEVVTTELKAAEAGASIQVWSLLPSSRLPSLGDRRPVLL 606
QY 661 GPRGLLRSOREHGCKTPASTWASPRPLLNWSRKAQOGKGTGGCPQCAPVKPAPPSLKH 720
DB 607 GPRGLLRSOREHGCKTPASTWASPRPLLNWSRKAQOGKGTGGCPQCAPVKPAPPSLKH 666
QY 721 GMP 723
DB 667 GMP 669

RESULT 5
US-10-697-266-4
; Sequence 4, Application US/10697266
; Publication No. US20040067522A1
; GENERAL INFORMATION:
; APPLICANT: RUSCH, Douglas et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001099-CIP-DIV2
; CURRENT APPLICATION NUMBER: US/10/697,266
; CURRENT FILING DATE: 2003-10-31
; PRIOR APPLICATION NUMBER: 10/274,878
; PRIOR FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: 09/849,334
; PRIOR FILING DATE: 2001-05-07
; PRIOR APPLICATION NUMBER: 09/773,371
; PRIOR FILING DATE: 2001-02-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 634
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-697-266-4
```

```
Query Match 61.8%; Score 2331; DB 12; Length 634;
Best Local Similarity 67.0%; Pred. No. 2.7e-195;
Matches 473; Conservative 49; Mismatches 112; Indels 72; Gaps 6;

QY 1 MILEERPDCAGAGEESPRQISRRKPRKTRVSLRGRREGRLDVCAMRCGRVHRGDPVR 60
DB 1 MILEERPDCAGAGEESPRQISRRKPRKTRVSLRGRREGRLDVCAMRCGRVHRGDPVR 26
QY 61 VHVGPMPNQLHVGCDLSLTQIQCGQLQSRRAQTHQIQDKELQMTGAENLYRATSNRVR 120
```

```
DB 27 -----KGYGSFVQNPQQLQSHRARLHQIISKELMRMTGAENLYRATSNRVR 74
QY 121 ETVALELSYNSNLIQLKELEELSGGVDPCRHSSEAVTVMIPMLPLGLKETKELDOWSTPLK 180
DB 75 ETVALELSYNSNLIQLKELEELSGGVDPCRHSSEAVTVMIPMLPLGLKETKELDOWSTPLK 134
QY 181 ELISVHFGECDGASYEAREIREALRQAMRTSPRNSGLELLETAYYNQCFIDARPLTPAR 240
DB 135 ELISVHFGECDGASYEAREIREALRQAMRTSPRNSGLELLETAYYNQCFIDARPLTPAR 194
QY 241 SLGLFFHWYDSLTCGPAQORALAFKESVLFNIGALHTQIGARODRSCTEGARRAMEAFQ 300
DB 195 SPGLLFFHWYDSLTCGPAQORALAFKESVLFNIGALHTQIGARODRSCTEGARRAMEAFQ 254
QY 301 RAAGAFSLIRENFHSHAPSPDMSAASLCALEQLMMAQAQECVFEGLSPPASMAPQDCIAQL 360
DB 255 RAAGAFSLIRENFHSHAPSPDMSAASLCALEQLMMAQAQECVFEGLSPPASMAPQDCIAQL 314
QY 361 RLAGEAAQVAAEYRLVHRTMAQPPVHDYVPSWTALVHVKAAYFRSLAHYHVAMALCDGS 420
DB 315 RLAGEAAQVAAEYRLVHRTMAQPPVHDYVPSWTALVHVKAAYFRSLAHYHVAMALCDGS 374
QY 421 PATEGELPTHQVLOPFTSSKPRGPVLPQLEERRQLGKAHLKRAILGQBEALRLHALC 480
DB 375 PA-KGELARQGHVF-QPSTPHEPLGPTLPQHPEDRRKLAKAHLKRAILGQBEALRLHALC 432
QY 481 RVLREVDLLRAVISQTLQRSIAKYAELDREDDFCEAAEAPDIQPKTHQKPEARMPLRSQ 540
DB 433 RVLKRVLDLLQVVTQALERSIAKYSQLEERDDFCEATEAPDIQPKTHQTPTE----- 483
QY 541 KGPDIHRLGLPLSVFSAKNRWLVGPVHLTRGEGGFGTLRGDSPLVIAAVIPGSAQAAA 600
DB 484 -----GPLSVFSTKNRWQLVGPVHMTREGGFGTLRGDSPLVIAAVIPGSAQAAA 534
QY 601 GLKEGDYIVSVNGQPCRWWRHAEVVTTELKAAEAGASIQVWSLLPSSRLPSLGDRRPVLL 660
DB 535 GLKEGDYIVSVNGQPCRWWRHAEVVTTELKAAEAGASIQVWSLLPSSRLPSLGDRRPVLL 594
QY 661 GPRGLLRSOREHGCKTPASTWASPRPLLNWSRKAQOGKGTGGCPQPC 706
DB 595 W-----NQREGGFETPMPTTRPWPILGWSRKNKQGTGSHDPDC 634
```

```
RESULT 6
US-10-274-878-4
; Sequence 4, Application US/10274878
; Publication No. US20030049792A1
; GENERAL INFORMATION:
; APPLICANT: RUSCH, Douglas et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001099-CIP-DIV
; CURRENT APPLICATION NUMBER: US/10/274,878
; CURRENT FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: 09/849,334
; PRIOR FILING DATE: 2001-05-07
; PRIOR APPLICATION NUMBER: 09/773,371
; PRIOR FILING DATE: 2001-02-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 634
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-274-878-4
```

```
Query Match 61.8%; Score 2331; DB 14; Length 634;
Best Local Similarity 67.0%; Pred. No. 2.7e-195;
Matches 473; Conservative 49; Mismatches 112; Indels 72; Gaps 6;

QY 1 MILEERPDCAGAGEESPRQISRRKPRKTRVSLRGRREGRLDVCAMRCGRVHRGDPVR 60
```

Best Local Similarity 67.0%; Pred. No. 2.8e-195;
Matches 473; Conservative 49; Mismatches 112; Indels 72; Gaps 6;

```
QY      1 MILLEERPDGAGAGEEPRQLQISRRKPKTVSSILRGREGRLDVCWARGCVRHGEDPVR 60
Db      1 MILLEERPDGGTGEESRPDD-----GSIR-----                26

QY     61 VHVGPMPNPQIHAVGCDSLTLQICQGQLQSRAAQIHQQIDKELOMRTGAENLYRATSNNRVR 120
Db      1 MILLEERPDGGTGEESRPDD-----GSIR-----                26

QY    121 ETVALELSYNLNQLLKBELSGSDVDGRHGSEAAVTVPMTPLGLKETKELDWSPTLK 180
Db      1 ETVALELSYNLNQLLKBELSLSTVDVDPEGEGITIPMTPLGLKETKELDWTATPLK 134

QY    181 ELISVFHGDGAYEAIELEALROAMRTSPNESGLELLLTAYNQCFLDFARFLTTPAR 240
Db    135 ELISEFHGDGTGFETIELEDLRQATKTPSDEAGLDLLAAYSQLCDFLARFFPSPR 194

QY    241 SLGLFHWYDLSLTGVPAQRALAFKXSVLFNIGALHTQIGARQDRSCITEGARAMEAFQ 300
Db    195 SPGLLFWYDLSLTGVPAQRALAFKXSVLFNIGALHTQIGARQDCSCITEGNHAEAFQ 254

QY    301 RAAGATSLRENESHAPSDDMSAASICALBOLMAQAQECVFGSLGEPAPMAQODCLAQ 360
Db    255 RAAGAFRLRENFSHAPSDDMSAASLSMLEQLMIAQAQECIFKGLLLPASATPDICPDQL 314

QY    361 RLQAQEAQAQAAVEVLVHRTHMAQPPVHDYVPVSNTALVHVKAEFERSLAHVHVAWLCDGS 420
Db    315 QLAQEAQAQAVATEVGLVHRAWAQPPVDYDPASWTNLIAHVKAHF'CALAHVHAAWALCESH 374

QY    421 PATYEGELPTHQEVFLQP'TSSKPRGVPVLPOELEERRQLGKAHLKRAILGOEEALRLHALC 480
Db    375 PA-KGELAQRHFV-QPSHPHEPLGPTLQHPPEDRRKLAKAHLKRAILGOEEALRLHTLC 432

QY    481 RVLRVEDLLRAVISQTOLRSKAYABLDREDDCEAAEAPDIOPKTHQKPPEARMPRLSQG 540
Db    433 RVLRKYVDLLQVVVTQALRSLSKAYSQLEREDDDFEATEAPDIQPKTHQTEP----- 483

QY    541 KGPDIHRGLPLSVFSAKNWRRLVGPHLTRGGGFGTLTRGDSVPLIAAVIPGSAQAAA 600
Db    484 -----GPLSVFTKNRWLVGPVHMTRGGGFGTLTRGDSVPLIAAVIPGGAESA 534

QY    601 GLKEGYIVSYNGQPCRWHRHAEVWTELKAAGAGAGASLOWVSLLPSSRLPSLGRDRPVLL 660
Db    535 GLKEGYIVSYNGQPCCKWKHLEVVTQLRSMBEGYSLVQVSLLPSPPEPGTGFRRAALL 594

QY    661 GPRGLLRSRHHGCKTPTASTWASPRPLLNSRKAAQOGKTGCQPQC 706
Db    595 W-----NQRECGETEMPTRTTRPWPILGWSRNKGKGTGSHDPDC 634
```

RESULT 8
US-10-408-765A-2465
; Sequence 2465, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408, 765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2465
; LENGTH: 311

```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-2465

Query Match      38.2%; Score 1442.5; DB 16; Length 311;
Best Local Similarity 91.6%; Pred. No. 9.3e-118;
Matches 282; Conservative 0; Mismatches 1; Indels 25; Gaps 1;

QY 380 MAQPPVHDYVPSWTALVHVHKAERYFRSLAHVHVAMALCDGS----- 420
Db 1 MAQPPVHDYVPSWTALVHVHKAERYFRSLAHVHVAMALCDGSRECPHPLMVLPRPPRAGS 60

QY 421 -----PATEGELPHEQVFLQPTTSSKPRGVPVLPQELERRQLGKAHLKRAILGQEEAL 474
Db 61 QPLCPAATEGELPHEQVFLQPTTSSKPRGVPVLPQELERRQLGKAHLKRAILGQEEAL 120

QY 475 RLHALCRVLRVDLLRAVISQTLQSLAKYAEALDREDDFCEAAAPDIQPKTHQKPEARM 534
Db 121 RLHALCRVLRVDLLRAVISQTLQSLAKYAEALDREDDFCEAAAPDIQPKTHQKPEARM 180

QY 535 PRLSQGKGPDIPIHRLGPLSVFSAKNRWRLVGPVHLTRGEGGFGLTIRGDSPLVLIAAVPIG 594
Db 181 PRLSQGKGPDIPIHRLGPLSVFSAKNRWRLVGPVHLTRGEGGFGLTIRGDSPLVLIAAVPIG 240

QY 595 SQAAAAGLKEGDYIVSVNGQPCRWWRHAEVVTTELKAAAGAGASLQVWSLLPSSRLPSLGD 654
Db 241 SQAAAAGLKEGDYIVSVNGQPCRWWRHAEVVTTELKAAAGAGASLQVWSLLPSSRLPSLGD 300

QY 655 RRPVLLGP 662
Db 301 RRPVLLGP 308

RESULT 9
US-09-925-301-1016
; Sequence 1016, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1016
; LENGTH: 257
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-301-1016

Query Match      33.9%; Score 1278; DB 9; Length 257;
Best Local Similarity 99.2%; Pred. No. 1.9e-103;
Matches 245; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 459 GKAKHLKRAILGQEEALRLHALCRVLRVDLLRAVISQTLQSLAKYAEALDREDDFCEAAE 518
Db 8 GKAKHLKRAILGQEEALRLHALCRVLRVDLLRAVISQTLQSLAKYAEALDREDDFCEAAE 67

QY 519 APDIQPKTHQKPEARMPLRSQKGPDIPIHRLGPLSVFSAKNRWRLVGPVHLTRGEGGFG 578
Db 68 APDIQPKTHQKPEARMPLRSQKGPDIPIHRLGPLSVFSAKNRWRLVGPVHLTRGEGGFG 127

QY 579 TLRGDSPVLIIAIVIPGSOAAAAGLKEGDYIVSVNGQPCRWWRHAEVVTTELKAAAGAGASL 638
Db 128 TLRGDSPVLIIAIVIPGSOAAAAGLKEGDYIVSVNGQPCRWWRHAEVVTTELKAAAGAGASL 187

QY 639 QVWSLLPSSRLPSLGDRLRPVLLGPRGLLRSQREHGCKTPTASTWASPRLLNWSRKAQOQK 698
Db 188 QVWSLLPSSRLPSLGDRLRPVLLGPRGLLRSQREHGCKTPTASTWASPRLLNWSRKAQOQK 247

QY 699 TGGCPQP 705
Db 248 TGGCPSP 254

RESULT 11
US-10-023-896-87
; Sequence 87, Application US/10023896
; Publication No. US20030027776A1
; GENERAL INFORMATION:
; APPLICANT: Victor Roschke
; TITLE OF INVENTION: 29 Human Cancer Associated Proteins
; FILE REFERENCE: PA004PI
; CURRENT APPLICATION NUMBER: US/10/023,896
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: unassigned
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: PCT/US00/23794
; PRIOR FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/152,296
```

```
/ PRIOR FILING DATE: 1999-09-03
/ PRIOR APPLICATION NUMBER: 60/158,003
/ PRIOR FILING DATE: 1999-10-06
/ NUMBER OF SEQ ID NOS: 138
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 87
/ LENGTH: 257
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-023-896-87

Query Match      33.9%; Score 1278; DB 14; Length 257;
Best Local Similarity 99.2%; Pred. No. 1.9e-103;
Matches 245; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 459 GKAKLRAILGQBEALRLHALCRVLREVDLLRAVISQTLQRSIAKYAELDRDDFCEAAE 518
Db 8 GKAKLRAILGQBEALRLHALCRVLREVDLLRAVISQTLQRSIAKYAELDRDDFCEAAE 67

QY 519 APDIQKTHQKPEARMRLSOGKGPDI FHLGLPLSVFSKAKNRRLVGPVHLTRGEGFGL 578
Db 68 APDIQKTHQKPEARMRLSOGKGPDI FHLGLPLSVFSKAKNRRLVGPVHLTRGEGFGL 127

QY 579 TLRGDSPVLIAAIPGSAAGLKEGDYIVSVNGQPCRWRAHAEVVTTELKAAAGEAGASL 638
Db 128 TLRGDSPVLIAAIPGSAAGLKEGDYIVSVNGQPCRWRAHAEVVTTELKAAAGEAGASL 187

QY 639 QVSVLLPSSRLPSLGRDRPVLLGPRGLLRQSRHGCKTTPASTWASPRLLNWSRKAQOQK 698
Db 188 QVSVLLPSSRLPSLGRDRPVLLGPRGLLRQSRHGCKTTPASTWASPRLLNWSRKAQOQK 247

QY 699 TGSCPOP 705
Db 248 TGSCPOP 254

RESULT 12
US-10-108-260A-4840
/ Sequence 4840, Application US/10108260A
/ Publication No. US20040005560A1
/ GENERAL INFORMATION:
/ APPLICANT: HELIX RESEARCH INSTITUTE
/ FILE REFERENCE: H1-A0106
/ CURRENT APPLICATION NUMBER: US/10/108,260A
/ CURRENT FILING DATE: 2002-03-27
/ NUMBER OF SEQ ID NOS: 5458
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 4840
/ LENGTH: 686
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-108-260A-4840

Query Match      33.3%; Score 1256; DB 15; Length 686;
Best Local Similarity 45.9%; Pred. No. 7.3e-101;
Matches 266; Conservative 107; Mismatches 196; Indels 10; Gaps 5;

QY 74 GDSITQIQGQIQSRRAQHQIQDKELQVTRGAENLYRATSNRVRVETVALELSYVNSN 133
Db 24 GCNPLAQTRGSKLQNGRAALNQILKAMRMRTGAENLLKVATNSKVREQVRLSESPVNS 83

QY 134 LQMLELEELSGVDPGRHSGEAVTVPMPLGLKETKELDWSITPLKELISVHFGEAGAS 193
Db 84 LQMLELEELSGVDPGRHSGEAVTVPMPLGLKETKELDWSITPLKELISVHFGEAGAS 143

QY 194 YEAEIRELEALQAMTPTSRNNEGLELLTAYYNOLCFDLARFLTPARSLGLPEHWDLSLT 253
Db 144 YEAEIRELEALQAMTPTSRNNEGLELLTAYYNOLCFDLARFLTPARSLGLPEHWDLSLT 203

QY 254 GYPAQORALAFKSGSVLFINIGALHTQIGARQSCRTGEGARRAMEAFQRAAGAFSLRENF 313
Db 204 GYPAQORALAFKSGSVLFINIGALHTQIGARQSCRTGEGARRAMEAFQRAAGAFSLRENF 263
```

```
QY 314 SHAPSPDMSAASICALQELMMAQAQECVFEGLSPASPAPQDCLAQRLRAQEAQAQVAAEY 373
Db 264 THTPSDYMSFAMLSVLVQKMLAQQAQESVFEKI SLPG--IRNEFFMLVKVQAQAQVAAEY 321

QY 374 RLVRHTMAQPPVHDYVVSWTALVHVKAQYFRSLAHYHVMALCDG--SPATEGELDTHE 431
Db 322 QQLHAAMSQAQPVKENIPYSWASLACVKAHHAALAHYFTAILLIDHQVKPGTDLID--HQ 378

QY 432 QVFLQPTSSKPRG--PVLIP-QELEBRRRLGKAHLKRAILGQBEALRLHALCRVLREVDL 488
Db 379 EKCLSQLYDHMEPEGLTFLATLKNQDQRRQLGKSLRRAMAHHEESVREASCLCKLSIEV 438

QY 489 LRAVISQTLQRSIAKYAELDRDDFCEAAEAPDIQKTHQKPEARMRLSOGKGPDI FHR 548
Db 439 LQKVLCAAGERSRLTYAQHQEEDDLNLIDAPSVAQTEQVDIILPQFSKLTVDFFQK 498

QY 549 LGPLSVFSKAKNRRLVGPVHLTRGEGFGLTRGDSPLVIAAIPGSAAGLKEGDYIVSV 608
Db 499 LGPLSVFSKAKNRRLVGPVHLTRGEGFGLTRGDSPLVIAAIPGSAAGLKEGDYIVSV 558

QY 609 VSVNGQPCRWRAHAEVVTTELKAAAGEAGASLQVSVLLPSS 647
Db 559 VSIQLVDCNKLITLSEVNMKLLKSGEDEIEKMKVSLDST 597

RESULT 13
US-10-092-900A-352
/ Sequence 352, Application US/10092900A
/ Publication No. US20040043382A1
/ GENERAL INFORMATION:
/ APPLICANT: Padigaru, Muralidhara
/ APPLICANT: Spyttek, Kimberly A.
/ APPLICANT: Shenoy, Suresh G.
/ APPLICANT: Taupier Jr., Raymond J.
/ APPLICANT: Pena, Carol E.A.
/ APPLICANT: Li, Li
/ APPLICANT: Zerhusen, Bryan D.
/ APPLICANT: Gusev, Vladimir Y.
/ APPLICANT: Ji, Weizhen
/ APPLICANT: Gorman, Linda
/ APPLICANT: Miller, Charles E.
/ APPLICANT: Kekuda, Ramesh
/ APPLICANT: Patturajan, Meera
/ APPLICANT: Gangolli, Esha A.
/ APPLICANT: Vernet, Corine A.M.
/ APPLICANT: Guo, Xiaojia Sasha
/ APPLICANT: Tchernev, Velizar T.
/ APPLICANT: Fernandes, Elma R.
/ APPLICANT: Casman, Stacie J.
/ APPLICANT: Malyankar, Uriel M.
/ APPLICANT: Gerlach, Valerie
/ APPLICANT: Liu, Yi
/ APPLICANT: Anderson, David W.
/ APPLICANT: Spaderna, Steven K.
/ APPLICANT: Catterton, Elina
/ APPLICANT: Leite, Mario W.
/ APPLICANT: Zhong, Haihong
/ APPLICANT: Alsobrook, John P.
/ APPLICANT: Lepley, Denise M.
/ APPLICANT: Rieger, Daniel K.
/ APPLICANT: Burgess, Catherine E.
/ TITLE OF INVENTION: No. US20040043382A1el Proteins and Nucleic Acids Encoding Same
/ FILE REFERENCE: 21402-290C
/ CURRENT APPLICATION NUMBER: US/10/092,900A
/ CURRENT FILING DATE: 2002-03-07
/ PRIOR APPLICATION NUMBER: USSN 60/274,322
/ PRIOR FILING DATE: 2001-03-08
/ PRIOR APPLICATION NUMBER: USSN 60/283,675
/ PRIOR FILING DATE: 2001-04-13
/ PRIOR APPLICATION NUMBER: USSN 60/338,092
/ PRIOR FILING DATE: 2001-12-03
/ PRIOR APPLICATION NUMBER: USSN 60/274,281
```

```
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: USN 60/274,191
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: USN 60/325,681
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: USN 60/304,354
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: USN 60/279,995
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: USN 60/294,899
; PRIOR FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: USN 60/287,424
; PRIOR FILING DATE: 2001-04-30
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 768
; SEQ ID NO 352
; LENGTH: 611
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-092-900A-352

Query Match      33.3%; Score 1255; DB 12; Length 611;
Best Local Similarity 45.9%; Pred. No. 7.5e-101;
Matches 266; Conservative 107; Mismatches 196; Indels 10; Gaps 5;

QY      74 GCDLSITQICGQLOSRRAQIHQOIDKELQMRGTGAENLYRATSNRVRETVALELSVNSN 133
DB      24 GCNPLAQTGRSKLQNRALNQILKAVMRGTGAENLLKVATNSKVRQVRLSEFVNSD 83

QY      134 LQLKEELEELSGGVDPRGHGSEAVTPMIPGLKETKELDWSTPKELISVHFGEDEGAS 193
DB      84 LQMLKEELEGLNISGVVQNTTEEAFTIPLIPLGLKETKDQVDFAVVLKDFILEHYSEDGYL 143

QY      194 YEAREIRELEALQAMRTSPRNESGLELLTAYYNQICFLDARPLTPARSIGLGFHHYDSLT 253
DB      144 YEDEIADLMDLRQACRTSPRDAGVELLMTYFIQIGFVESRFPPTQMGLLFTWYDSLT 203

QY      254 GYPAQORALAFKGSVLFNIGALHTQIGARODRSCTEGARRAMEAFQRAAGAFSLLENF 313
DB      204 GYFVSQONLLEKASVLFNTGALYTIQIGTRCDRQTCAGLESIDAIDAFQRAAGVNLVKDTF 263

QY      314 SHAPSPDMSAASLCALQELQMAQAQECVFEGSLPPASMAPQCLALQRLAQAAQVAAEY 373
DB      264 THTPSYDMSAPMLSVLVKMLAQAESVFEKISLPG--IRNEFFMLVKVQAQAAKVGEVY 321

QY      374 RLAVHTMAQPPVHDVVPVSWTALVHKAEYFSLAHYHVMALCDG--SPATEGELPTHE 431
DB      322 QQLHAAMSQAPVKENIPYSWASLACVKAHHAALAHYFTAILLIDHQVKPGDLDL--HQ 378

QY      432 QVFLQPTSSKPRG--PVLP--QELERRQLGKAHLKRAILGOEALRLHALCRVLRVDL 488
DB      379 EKCLSQLYDHMPEGLTPLATLKNQDQRRQLGKSHLRRAHMAHHEESVREASLCKLRSIEV 438

QY      489 LRAVISQTLQRLSLAKYAEILDREDDFCEAAEAPDIQPKTHQKPEARMPLRSQKGGPDIFHR 548
DB      439 LQKVLCAAQERSRLTYAQHQEEDLLNLIDAPSVVAKTEQEVDDIILPQFSKLTVTDFFOK 498

QY      549 LGPLSVFSAKNRWLRVGLTRGGFGLTRGSPVLIAAVIPGSAQAAAGLKGEDYI 608
DB      499 LGPLSVFSANKRWTPPRIRFTAEGDLGFTLRGNAPVQVHFDPYCSASVAGAREGDYI 558

QY      609 VSVNGQPCRWRRHAEEVVTTELKAAAGASQVSVLLPSS 647
DB      559 VSIQLVDCKWLTLSEVMKLLKSGFGEDEIEMKVVSLLDST 597

RESULT 14
US-10-239-607-53
; Sequence 53, Application US/10239607
; Publication No. US20030219761A1
; GENERAL INFORMATION:
; APPLICANT: Saatcioglu, Fahri
; TITLE OF INVENTION: No. US20030219761A1el Prostate-Specific or

; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: USN 60/274,191
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: USN 60/325,681
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: USN 60/304,354
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: USN 60/279,995
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: USN 60/294,899
; PRIOR FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: USN 60/287,424
; PRIOR FILING DATE: 2001-04-30
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 768
; SEQ ID NO 352
; LENGTH: 611
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-092-900A-352

Query Match      33.3%; Score 1255; DB 12; Length 611;
Best Local Similarity 45.9%; Pred. No. 7.5e-101;
Matches 266; Conservative 107; Mismatches 196; Indels 10; Gaps 5;

QY      74 GCDLSITQICGQLOSRRAQIHQOIDKELQMRGTGAENLYRATSNRVRETVALELSVNSN 133
DB      24 GCNPLAQTGRSKLQNRALNQILKAVMRGTGAENLLKVATNSKVRQVRLSEFVNSD 83

QY      134 LQLKEELEELSGGVDPRGHGSEAVTPMIPGLKETKELDWSTPKELISVHFGEDEGAS 193
DB      84 LQMLKEELEGLNISGVVQNTTEEAFTIPLIPLGLKETKDQVDFAVVLKDFILEHYSEDGYL 143

QY      194 YEAREIRELEALQAMRTSPRNESGLELLTAYYNQICFLDARPLTPARSIGLGFHHYDSLT 253
DB      144 YEDEIADLMDLRQACRTSPRDAGVELLMTYFIQIGFVESRFPPTQMGLLFTWYDSLT 203

QY      254 GYPAQORALAFKGSVLFNIGALHTQIGARODRSCTEGARRAMEAFQRAAGAFSLLENF 313
DB      204 GYFVSQONLLEKASVLFNTGALYTIQIGTRCDRQTCAGLESIDAIDAFQRAAGVNLVKDTF 263

QY      314 SHAPSPDMSAASLCALQELQMAQAQECVFEGSLPPASMAPQCLALQRLAQAAQVAAEY 373
DB      264 THTPSYDMSAPMLSVLVKMLAQAESVFEKISLPG--IRNEFFMLVKVQAQAAKVGEVY 321

QY      374 RLAVHTMAQPPVHDVVPVSWTALVHKAEYFSLAHYHVMALCDG--SPATEGELPTHE 431
DB      322 QQLHAAMSQAPVKENIPYSWASLACVKAHHAALAHYFTAILLIDHQVKPGDLDL--HQ 378

QY      432 QVFLQPTSSKPRG--PVLP--QELERRQLGKAHLKRAILGOEALRLHALCRVLRVDL 488
DB      379 EKCLSQLYDHMPEGLTPLATLKNQDQRRQLGKSHLRRAHMAHHEESVREASLCKLRSIEV 438

QY      489 LRAVISQTLQRLSLAKYAEILDREDDFCEAAEAPDIQPKTHQKPEARMPLRSQKGGPDIFHR 548
DB      439 LQKVLCAAQERSRLTYAQHQEEDLLNLIDAPSVVAKTEQEVDDIILPQFSKLTVTDFFOK 498

QY      549 LGPLSVFSAKNRWLRVGLTRGGFGLTRGSPVLIAAVIPGSAQAAAGLKGEDYI 608
DB      499 LGPLSVFSANKRWTPPRIRFTAEGDLGFTLRGNAPVQVHFDPYCSASVAGAREGDYI 558

QY      609 VSVNGQPCRWRRHAEEVVTTELKAAAGASQVSVLLPSS 647
DB      559 VSIQLVDCKWLTLSEVMKLLKSGFGEDEIEMKVVSLLDST 597

RESULT 15
US-09-895-040A-3
; Sequence 3, Application US/09895040A
; Patent No. US20020123474A1
; GENERAL INFORMATION:
; APPLICANT: Shannon, Mark
; TITLE OF INVENTION: HUMAN GTP-RHO BINDING PROTEIN 2
; FILE REFERENCE: AEWICA-11
; CURRENT APPLICATION NUMBER: US/09/895,040A
; CURRENT FILING DATE: 2001-06-29
```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 13, 2004, 08:26:41 ; Search time 25 Seconds
(without alignments)
1505.869 Million cell updates/sec

Title: US-10-697-266-2

Perfect score: 3774

Sequence: 1 MILEERPDCGAGEBSPRLQ.....QPCAPVKPAPPSLKHGPWP 723

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	3419	90.6	695	1 RHNI_HUMAN	Q8tcx5 mus sapien
2	2331	61.8	643	1 RHNI_MOUSE	Q61085 mus musculu
3	284.5	7.5	861	1 PQS8_CAEEL	P34552 caenorhabdi
4	274.5	7.3	868	1 PD61_HUMAN	Q8wum4 homo sapien
5	266.5	7.1	869	1 PD61_MOUSE	Q9wu78 mus musculu
6	261.5	6.9	867	1 PD61_XENLA	Q9w6c5 xenopus lae
7	202	5.4	844	1 BR01_YEAST	P48582 saccharomyc
8	156.5	4.1	942	1 PKL1_HUMAN	Q16512 homo sapien
9	146.5	3.9	946	1 PKL1_RAT	Q63433 rattus norv
10	146	3.9	1017	1 EM11_MOUSE	Q99x41 mus musculu
11	145.5	3.9	1522	1 ARHB_HUMAN	Q15085 homo sapien
12	138	3.7	910	1 US1C_MOUSE	Q9es64 mus musculu
13	136.5	3.6	1527	1 ARHB_RAT	Q9es67 rattus norv
14	135.5	3.6	1016	1 EM11_HUMAN	Q9y5c2 homo sapien
15	132	3.5	1544	1 ARHC_HUMAN	Q9nzn5 homo sapien
16	131.5	3.5	838	1 KFC2_HUMAN	Q96ac6 homo sapien
17	131.5	3.5	2564	1 SPQC_HUMAN	Q9h254 homo sapien
18	130	3.4	1387	1 RGS_C RAT	Q08774 rattus norv
19	129	3.4	1205	1 SMC MYCTU	Q10970 mycobacteri
20	128.5	3.4	1455	1 AIPI_HUMAN	Q86ul8 homo sapien
21	128.5	3.4	3674	1 SPKR_HUMAN	Q9nrc6 homo sapien
22	128	3.4	1543	1 ARHC_MOUSE	Q8r4h2 mus musculu
23	127.5	3.4	4684	1 PLE1_HUMAN	Q15149 homo sapien
24	126	3.3	4687	1 PLE1_RAT	P30427 rattus norv
25	124	3.3	701	1 YAB5_SCHPO	Q9wyq1 mus musculu
26	124	3.3	1275	1 AIPI_MOUSE	Q9j155 cricetus
27	124	3.3	4473	1 PLE1_CRIGR	Q16363 homo sapien
28	123.5	3.3	1816	1 LMA4_HUMAN	Q88382 rattus norv
29	121.5	3.2	1277	1 AIPI_RAT	Q9zix6 thermus cal
30	121	3.2	817	1 MUT5_THECA	Q14924 homo sapien
31	121	3.2	1447	1 RGS_C_HUMAN	Q56239 thermus the
32	119.5	3.2	818	1 MUT5_THETH	P48725 mus musculu
33	119	3.2	1920	1 PCT2_MOUSE	

34	119	3.2	5327	1 MACF_MOUSE	Q9qxx0 mus musculu
35	118	3.1	330	1 RIL_CHICK	Q9pw72 gallus gall
36	118	3.1	651	1 PRIA_MYCLE	Q9ccq3 mycobacteri
37	118	3.1	1445	1 PTPG_HUMAN	P23470 homo sapien
38	117	3.1	668	1 FAU_DROME	Q9vqx3 drosophila
39	116.5	3.1	531	1 GL1A_ORYZA	F47997 oryza sativ
40	116.5	3.1	1739	1 CHD2_HUMAN	O14647 homo sapien
41	116	3.1	946	1 RHG4_HUMAN	P98171 homo sapien
42	116	3.1	1238	1 SBCC_HOCCA	O68032 rhodobacter
43	116	3.1	2442	1 CEP2_HUMAN	Q9bw73 homo sapien
44	115.5	3.1	655	1 PRIA_MYCTU	P71670 mycobacteri
45	115.5	3.1	933	1 ZO3_HUMAN	O95049 homo sapien

ALIGNMENTS

RESULT 1
RHNI_HUMAN
ID RHNI_HUMAN STANDARD; PRT; 695 AA.
AC Q8TCX5;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Rhophilin 1.
GN RHPN1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Burbelo P.D.;
RT "Structure and function of rhophilin homologs.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Binds specifically to GTP-Rho. May serve as a target for
CC Rho. It has no enzymatic activity. It may interact with some
CC cytoskeletal component upon Rho binding or relay a Rho signal to
CC other molecules (By similarity).
CC -!- SIMILARITY: Contains 1 BR01 domain.
CC -!- SIMILARITY: Contains 1 PDZ/DHR domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AY082588; AAL89809.1; -
CC Genew; HGNC:19973; RHPN1.
CC InterPro; IPR004328; BR01.
CC InterPro; IPR001478; PDZ.
CC InterPro; IPR000861; REM_repeat.
CC Pfam; PF03097; BR01; 1.
CC Pfam; PF02185; HR1; 1.
CC Pfam; PF00595; PDZ; 1.
CC SMART; SM00074; HR1; 1.
CC SMART; SM00228; PDZ; 1.
CC PROSITE; PS50106; PDZ; 1.
CC Q15149 homo sapien
CC FT DOMAIN 35 107 HR1.
CC FT DOMAIN 108 260 BR01.
CC FT DOMAIN 538 617 PDZ.
CC SQ SEQUENCE 695 AA; 76292 MW; 89E4DF6806761452 CRC64;

Query Match 90.6%; Score 3419; DB 1; Length 695;

Best Local Similarity 89.4%; Pred. No. 1.1e-218;

Matches 669; Conservative 0; Mismatches 1; Indels 78; Gaps 2;

QY 1 MILEERPDCGAGEBSPRLQISRRXPRKTRVSSLRGRREGLRDVCAMRGCRVHRGEDPVR 60
|||||

```
Db 1 MILERPDCGAGABESPRLO----- 20
Qy 61 VHVGMNPNQLHVGCDLSLTQICGQLOSRRAQIHQOIDKELQMTGAENLYRATSNRVR 120
Db 21 -----GCDLSLTQICGQLOSRRAQIHQOIDKELQMTGAENLYRATSNRVR 67
Qy 121 ETVALELSYVNSNLQLLKEELEELSGGVDPRHGSEAVTVPMIPLGLKETKELDWSTPLK 180
Db 68 ETVALELSYVNSNLQLLKEELEELSGGVDPRHGSEAVTVPMIPLGLKETKELDWSTPLK 127
Qy 181 ELISVHFGEDGASVEAEIRELEALRQAMRTPSRNSGLELLTAYYNQLCFDARFLTPAR 240
Db 128 ELISVHFGEDGASVEAEIRELEALRQAMRTPSRNSGLELLTAYYNQLCFDARFLTPAR 187
Qy 241 SLGLFFHWYDSLTVGPAQORALAFKGVLFNIGALHTQIGARQDRSCTEGARRAMEAFQ 300
Db 188 SLGLFFHWYDSLTVGPAQORALAFKGVLFNIGALHTQIGARQDRSCTEGARRAMEAFQ 247
Qy 301 RAAGAFSLLRNFHSHAPSPDMSAASLCALQLEQLMMAQAQCEVPEGLSPSPASMAPQDCLAOL 360
Db 248 RAAGAFSLLRNFHSHAPSPDMSAASLCALQLEQLMMAQAQCEVPEGLSPSPASMAPQDCLAOL 307
Qy 361 RLAEAAQVAAAEYRLVHRTMAQPPVHDVVPVSWTALVHVKAIEYFRSLAHYHVAMALCDGS 420
Db 308 RLAEAAQVAAAEYRLVHRTMAQPPVHDVVPVSWTALVHVKAIEYFRSLAHYHVAMALCDGS 367
Qy 421 -----PATEGELPTHEQVFLQPTSSKPRGPVLPQLEER 455
Db 368 RECPLHFMVLPRPPRAGSQPLCPAATGELPTHEQVFLQPTSSKPRGPVLPQLEER 427
Qy 456 RQLGAHLKRALIGQEEALRHLCRLVRLVDLLRAVLSQTLQRLAKYAELEDREDDFCE 515
Db 428 RQLGAHLKRALIGQEEALRHLCRLVRLVDLLRAVLSQTLQRLAKYAELEDREDDFCE 487
Qy 516 AAAPDIQPKTHQKPEARNPRLSQKGPDI FRLGLPVSFAKNRWLVGPVHLTRGEGG 575
Db 488 AAAPDIQPKTHQKPEARNPRLSQKGPDI FRLGLPVSFAKNRWLVGPVHLTRGEGG 547
Qy 576 FGLTLRGSPVLIAAVIPGSAAGLKEGDYIVSVNGQPCRWRRHAEVVTCLKAAGBAG 635
Db 548 FGLTLRGSPVLIAAVIPGSAAGLKEGDYIVSVNGQPCRWRRHAEVVTCLKAAGBAG 607
Qy 636 ASLVQVLLPSRLPSLGRDRPVLLGPRGLRSQREHGCKTPASTWASPRPLLNWSRKAQ 695
Db 608 ASLVQVLLPSRLPSLGRDRPVLLGPRGLRSQREHGCKTPASTWASPRPLLNWSRKAQ 667
Qy 696 QGKTGGCPQCAPVKPAPPSSLKHPGW 723
Db 668 QGKTGGCPQCAPVKPAPPSSLKHPGW 695
```

RESULT 2

```
RHNI_MOUSE
ID RHNI_MOUSE STANDARD; PRT; 643 AA.
AC Q61085;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Rhophilin 1 (GTP-Rho binding protein 1).
GN RHPN1 OR GRBP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain, and Embryo;
RX MEDLINE=96165390; PubMed=8571126;
RA Watanabe G., Saito Y., Madaule P., Ishizaki T., Fujisawa K., Morii N.,
RA Mukai H., Ono Y., Kakizuka A., Narumiya S.;
RT "protein kinase N (PKN) and PKN-related protein rhophilin as targets
RT of small GTPase Rho."
RL Science 271:645-648(1996).
```

```
CC -!- FUNCTION: Binds specifically to GTP-Rho. May serve as a target for
CC Rho. It has no enzymatic activity. It may interact with some
CC cytoskeletal component upon Rho binding or relay a Rho signal to
CC other molecules.
CC -!- TISSUE SPECIFICITY: Highly expressed in testis.
CC -!- SIMILARITY: Contains 1 BRO1 domain.
CC -!- SIMILARITY: Contains 1 PDZ/DHR domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U43194; AAC52388.1; -.
CC HSSP: P29476; IQAV.
CC MGD: MGI:1098783; Rbpnl.
CC GO: GO:0005083; F:small GTPase regulatory/interacting protein. .; IPI.
CC InterPro: IPR004328; BRO1.
CC InterPro: IPR001478; PDZ.
CC InterPro: IPR000861; REM_repeat.
CC Pfam: PF03097; BRO1; 1.
CC Pfam: PF02185; HRI; 1.
CC Pfam: PF00595; PDZ; 1.
CC SMART: SM00074; HRI; 1.
CC SMART: SM00228; PDZ; 1.
CC PROSITE: PS50106; PDZ; 1.
CC DOMAIN 42 114 HRI.
CC FT DOMAIN 115 267 BRO1.
CC FT DOMAIN 500 577 PDZ.
CC SEQUENCE 643 AA; 71288 MW; 5CD24E1417D8C6D8 CRC64;
CC
Qy Query Match 61.8%; Score 2331; DB 1; Length 643;
Qy Best Local Similarity 67.0%; Pred. No. 9.7e-147; Indels 72; Gaps 6;
Qy Matches 473; Conservative 49; Mismatches 112;
Db 1 MILERPDCGAGABESPRLOISRRKPRKTRVSSLSRREGRLDVCAMRCVHRGEDPVR 60
Db 1 MILERPDCGAGABESPRLOISRRKPRKTRVSSLSRREGRLDVCAMRCVHRGEDPVR 26
Qy 61 VHVGMNPNQLHVGCDLSLTQICGQLOSRRAQIHQOIDKELQMTGAENLYRATSNRVR 120
Db 27 -----KGYGSFVQNPQGLQSHRARLHQOISKELEMRGTCAENLYRATSNRVR 74
Qy 121 ETVALELSYVNSNLQLLKEELEELSGGVDPRHGSEAVTVPMIPLGLKETKELDWSTPLK 180
Db 75 ETVALELSYVNSNLQLLKEELEELSGGVDPRHGSEAVTVPMIPLGLKETKELDWSTPLK 134
Qy 181 ELISVHFGEDGASVEAEIRELEALRQAMRTPSRNSGLELLTAYYNQLCFDARFLTPAR 240
Db 135 ELISVHFGEDGASVEAEIRELEALRQAMRTPSRNSGLELLTAYYNQLCFDARFLTPAR 194
Qy 241 SLGLFFHWYDSLTVGPAQORALAFKGVLFNIGALHTQIGARQDRSCTEGARRAMEAFQ 300
Db 195 SPGLLFHWYDSLTVGPAQORALAFKGVLFNIGALHTQIGARQDRSCTEGARRAMEAFQ 254
Qy 301 RAAGAFSLLRNFHSHAPSPDMSAASLCALQLEQLMMAQAQCEVPEGLSPSPASMAPQDCLAOL 360
Db 255 RAAGAFSLLRNFHSHAPSPDMSAASLCALQLEQLMMAQAQCEVPEGLSPSPASMAPQDCLAOL 314
Qy 361 RLAEAAQVAAAEYRLVHRTMAQPPVHDVVPVSWTALVHVKAIEYFRSLAHYHVAMALCDGS 420
Db 315 RLAEAAQVAAAEYRLVHRTMAQPPVHDVVPVSWTALVHVKAIEYFRSLAHYHVAMALCDGS 374
Qy 421 PATEGELPTHEQVFLQPTSSKPRGPVLPQLEER 455
Db 375 PA-KELARQEHVP-QPSTPHEPLGTLPQHPEDRRKAKAHLKRALIGQEEALRHLCRLVRLVDLLRAVLSQTLQRLAKYAELEDREDDFCEAAEAPDIQPKTHQKPEARNPRLSQK 540
Qy 481 RVLRVDLLRAVLSQTLQRLAKYAELEDREDDFCEAAEAPDIQPKTHQKPEARNPRLSQK 540
```

Db 433 RVLKRVLDLQVVVVTQALRRSLAKYQSLEREDDFFATEAPDIQPKHOTPE----- 483
 QY 541 KCPDIFHRLGPLSVFSAXNRWRLVGVPHLTRGEGFGFLTRGDSPLVIAAVIPGSAAAA 600
 Db 484 -----GPLSVFSTKRWQLVGVPHVTRGEGFGFLTRGDSPLVIAAVIPGSAESA 534
 QY 601 GLKEGDIYVNVGQPCRWHRHAEVVTTELKAAAGASLQVSLPSRLPSIGDRRPVLL 660
 Db 535 GLKEGDIYVNVGQPCRWKHLVTVTLRSMGEEGVSLQWVSLPSPEPRGTGPRRAALL 594
 QY 661 GPRGLLRSOREHGCKTPASTWSPRLLNWSRKAQOGKGGCPQPC 706
 Db 595 W-----NORECGFETPEMTRTPWPILGWSRKNQKGTGSHPDPC 634

RESULT 3
 ID P058 CAEEL STANDARD; PRT; 861 AA.
 AC P34552; Q94159;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Protein pgn-58 (Protein YNK1).
 GN PQN-58 OR R10E12.1.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RX MEDLINE=98087404; PubMed=9427532;
 RA Che S., Weil M.M., Etkin L.D., Epstein H.F., Kuang J.;
 RT "Molecular cloning of a splice variant of Caenorhabditis elegans YNK1,
 a putative element in signal transduction."
 RL Biochim. Biophys. Acta 1354:231-240 (1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RX MEDLINE=94150718; PubMed=7906398;
 RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M., Coulson A.,
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Fraser A.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favell A., Fraser A.,
 RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
 RA Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,
 RA Latreille P., Lighthouse J., Lloyd C., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkeen R.,
 RA Sims M., Smaldon N., Smith A., Smith M., Sonhammer E., Staden R.,
 RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
 RA Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,
 RA Wohldman P.;
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 elegans."
 RL Nature 368:32-38 (1994).
 RN [3]
 RP REVISIONS.
 RA Durbin R.;
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Probably involved in signal transduction.
 CC -!- SIMILARITY: TO MAMMALIAN GRBP; SOME, TO YEAST BROL.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U73679; AAC67305.1; -
 CC EMBL; Z29561; CAA82667.2; -
 CC WormPep; R10E12.1; CE25075.
 CC InterPro; IPR004328; BROL.

DR Pfam; PF03097; BROL1; 1.
 FT DOMAIN 756 861 GLN-RICH.
 SQ SEQUENCE 861 AA; 96245 MW; F2F600623FF74965 CRC64;
 Query Match 7.5%; Score 284.5; DB 1; Length 861;
 Best Local Similarity 22.8%; Pred. No. 3.7e-11;
 Matches 132; Conservative 98; Mismatches 209; Indels 139; Gaps 22;
 QY 167 LKETELDKMTPLKELIS--VHFGEDGAGSYAE--IRELEAIR-QAMRTP-SRNESGLELL 221
 Db 11 LKSTNEVDLVKLPTSYIDNVNTSDNRSDDVAEAVQELNKLKSKACCQQLDHHQSALDVL 70
 QY 222 TAYYNOLCFDLARFLTPARSGLGFHHWYDSL-TGVPAQORAL-----AFEXKGSVFNIG 274
 Db 71 TRYDOLVAIENKIITISATQNVFVKNDAPFKGSLFSSRASLSLSDGSFERAAVLFNIG 130
 QY 275 ALHTOIGARQDRCTEGARRAMEAFORAAAFSLRE----NFSHAPSDDMSAASICAL 330
 Db 131 SLMSQIGAAQQFHTDDEIKVSAKLFOQSAGVFAFLRDVVLGMVQOEFTPLDMPDTILAALS 190
 QY 331 QLMMAQAQECVFEGLSPASMAPQDCLAQRLAQEAQAAVAAEYRLVHRTWAQPPVHDYVP 330
 Db 191 ALMTAQAEAIY-----IKGHEKMKATSMVKISAQVAEFYSEAKQMKSKDIVRGLWD 243
 QY 391 VSWTALVHVKAEVERSLAHVHVAMALCDGSPATEGELPTHEQVFLQPTTSKPRGPVLPQ 450
 Db 244 KWSALVSGKNLAYQALAQYHQS-EVCG-----
 QY 451 ELEERRQLGKAHLKRAILGQEEALRLHALCRVLRVDLLRAVISQ-----TLQSLAK- 503
 Db 271 ---EARQIGE-----QLSRLAESLKLFDTAQKYLPRDITGTWIDTPSVSKA 313
 QY 504 YALREDDFCEAAEAPDQPKTHQKPEAMPRLSQGKP-----DIEHRLGELSVEFA 557
 Db 314 HAAAKKDNDFIYHEKVSDFE-TLPTLPKAVIAKPTMQTPTMTPTSFDFMFAVLVPVQVHNA 372
 QY 558 KRW-----RLVGPVHLTRGEGGGLTLRGDSPVLIAAVIPGSAAGLKE 604
 Db 373 MQSYDARKAELVNMTVRMEATQLMNG-----VLASLNLPAAL----- 411
 QY 605 GDYIVSVNGQPCRWHRHAEVVTTELKAA--GEAGASLQVSLPSRLPSIGDRRPVLLGP 662
 Db 412 -DDVTSFETLP-----ESLKLKSAKLQNGSGSSEIMRLF--SELPTLYQRNEDILTE 460
 QY 663 RGLLSOREHGCKTPASTWSPRLLNWSRKAQOGKGTG 700
 Db 461 TSRILNEEKESDDTMRKQLGT-----KWTMRMSSEQLTG 493

RESULT 4
 ID PD61 HUMAN STANDARD; PRT; 868 AA.
 AC Q8WUM4; Q9BX86; Q9NUN0; Q9P2H2; Q9UKL5;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Programmed cell death 6-interacting protein (PDCD6-interacting
 DE protein) (ALG-2 interacting protein 1) (Hp95).
 GN PDCD6IP OR AIP1 OR KIAA1375.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=215339686; PubMed=11683497;
 RA Wu Y., Pan S., Che S., He G., Nelman-Gonzalez M., Weil M.M., Kuang J.;
 RT "Overexpression of Hp95 induces G1 phase arrest in confluent Hela
 cells."
 RL Differentiation 67:139-153 (2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Li H., Shioda T., Isselbacher K.J.;

RT "Molecular cloning of human ALG-2 interacting protein 1 (AIP1).";
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Kalnine N., Chen X., Rolfs A., Halleck A., Hines L., Eisenstein S.,
 RA Koundinya M., Raphael J., Moreira D., Kelley T., Labaer J., Lin Y.,
 RA Phelan M., Farmer A.,
 RA "Cloning of human full-length cDNAs in BD Creator(TM) system donor
 RT vector.";
 RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lymph;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Brakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Gricham J.J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RA "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [5]
 RP SEQUENCE OF 209-868 FROM N.A.
 RC TISSUE=Placenta;
 RA Isoqai T., Ota T., Hayaashi K., Sugiyama T., Otsuki T., Suzuki Y.,
 RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
 RA Watanabe S., Hosoi T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
 RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
 RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saio K.,
 RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahara K., Masuho Y.,
 RA Ninomiya K., Iwayanagi T.;
 RA "NEDO human cDNA sequencing project";
 RT Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE OF 323-868 FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=20181126; PubMed=10718198;
 RA Nagase T., Kikuno R., Ishikawa K.-I., Hirose M., Ohara O.;
 RA "Prediction of the coding sequences of unidentified human genes. XVI.
 RT The complete sequences of 150 new cDNA clones from brain which code
 RL for large proteins in vitro.";
 RN DNA Res. 7:65-73(2000).
 CC -!- FUNCTION: May play a role in the regulation of both apoptosis and
 CC cell proliferation.
 CC -!- SUBUNIT: Interacts with PDCD6/ALG-2 and SH3BP1 (By similarity).
 CC -!- SIMILARITY: Contains 1 BRO1 domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. Usage by and for commercial
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL; AF349951; AAK20398.1; -
 CC EMBL; AF151793; AAF08220.1; -
 CC EMBL; BT007367; AAP36031.1; -
 CC EMBL; BC020066; AAH20066.1; -
 CC EMBL; AK002122; BAA32092.1; ALT_INIT.

DR EMBL; AB037796; BAA92613.1; -.
 DR Genew; HGNC:8766; PDCD6IP.
 DR MIM; 608074; -.
 DR InterPro; IPR004328; BRO1.
 DR Pfam; PF03097; BRO1; 1.
 KW Apoptosis.
 FT DOMAIN 1 167 BRO1.
 FT DOMAIN 717 860 PRO-RICH.
 FT CONFLICT 309 309 A -> T (IN REF. 2).
 FT CONFLICT 378 378 V -> I (IN REF. 5).
 FT CONFLICT 550 550 N -> S (IN REF. 1 AND 6).
 FT CONFLICT 580 580 M -> T (IN REF. 5).
 FT CONFLICT 730 730 S -> L (IN REF. 2).
 SQ SEQUENCE 868 AA; 96023 MW; 573588D1F612EC93 CRC64;
 Query Match 7.3%; Score 274.5; DB 1; Length 868;
 Best Local Similarity 27.2%; Pred. No. 1.7e-10;
 Matches 116; Conservative 57; Mismatches 152; Indels 101; Gaps 17;
 QY 163 IPGLKETKELDWSTPLKELISVHF---GEGASVEAEIRELEALRQAM--RTPSRNESC 217
 DB 5 ISVOLKKTSEVLLAKELVFIQOTYPSGGEQAOYCRAAEELSKLRRAVGRFLDKHEGA 64
 QY 218 LELLTAYYNQLCFLDARFLTPARSLGLFFHWYD-----SLTG--VPAQORALAFKGSYL 270
 DB 65 LETLLRYYPQICSIIBPKFPFSENQICLTFTWKDAPDKGSLFGSVKLALSLGYEKSCVL 124
 QY 271 ENIGALHTOIGARODSCTEGHARRAMEAFQRAAGFSLREN----FSHAPSMDMSAASL 326
 DB 125 FNCALLASQIAABQNLNDDEGLKIAKHQYQFAGFLHIKETVLSALSREPTVDISPTDV 184
 QY 327 CALEOLIMMAQACQECVFEGLSPSPASMAPQDCLAQLRAQEAQVAAYEYRLVHRTMAQPPVH 386
 DB 185 GTLSLIMLAQAQEVFF-----LKATRKMDKAIKLQANQ--AADY--FGDAFKOCQYK 234
 QY 387 DYVPVSVTALVHVKAERYFSLAHYHVAMALCDGSPATEGELTHFQVFLQPTSSKPRGS 446
 DB 235 DTLP-----KEVPVLAAKHCIM-----QANAAYHQSI-----262
 QY 447 VLPQLEBERRQLGKHLKRAILGQEARLRLHLCRVLEVDLLRAV-----TS 494
 DB 263 -----LAKQKK---FGESIALQHA-----AELIKTVASGYDEVVNVKDFS 301
 QY 495 QTLQSLAKYAELEDDREDDCEAAEAPDIQ-----PKTHQKPEARMPLRSQCKGPDIFHR 548
 DB 302 DKINEAL---AAAKDNDFIHYDRVPLDKLDLPICKAILVKSPTVNVPLSQ--KFTDLPEK 357
 QY 549 LGPLSV 554
 DB 358 MVPVSV 363
 RESULT 5
 PDC1_MOUSE
 ID PDC1_MOUSE STANDARD; PRT; 869 AA.
 AC Q9WU78; O88695; O89014; Q8BSL8; O8RH5; Q99LR3; Q9CZNR;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Programmed cell death 6 interacting protein (ALG-2 interacting protein
 DE X) (ALG-2 interacting protein 1) (E2F1-inducible protein) (Big2).
 GN PDCD6IP OR ALIX OR AIP1.
 OS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CC NCBT_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), AND INTERACTION WITH PDCD6.
 RC TISSUE=Brain;
 RX MEDLINE=99218669; PubMed=10200558;
 RA Missorten M., Nichols A., Rieger K., Sadoul R.;
 RT "Alix, a novel mouse protein undergoing calcium-dependent interaction
 with the apoptotic-linked-gene (ALG-2) protein.";

Cell Death Differ. 6:124-129(1999).
 [2]
 RN SEQUENCE FROM N.A. (ISOFORM 1).
 RP MEDLINE=99098896; PubMed=9880530;
 RX Vito P., Pellegrini L., Guet C., D'Adamo L.;
 RT "Cloning of Aip1, a novel protein that associates with the apoptosis-
 linked gene ALG-2 in a Ca2+-dependent reaction.";
 RL J. Biol. Chem. 274:1533-1540(1999).
 [3]
 RN SEQUENCE FROM N.A. (ISOFORM 3).
 RP STRAIN=C57BL/6J; TISSUE=Forelimb;
 RX MEDLINE=22354683; PubMed=12466851;
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
 RA Nikaiko I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
 RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
 RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
 RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
 RA Dalia E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
 RA Grimond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
 RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
 RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
 RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
 RA Ravasi T., Reed J.C., Reid J., Reid J., Ring B.Z., Ringwald M.,
 RA Sandelin A., Schneider C., Sempile C.A., Setou M., Shimada K.,
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
 RA Verardo R., Wagner L., Wahlstedt C., Wang Y., Watanabe Y., Wells C.,
 RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
 RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
 RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
 RA Birney E., Hayashizaki Y.;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 60,770 full-length cDNAs.";
 RT Nature 420:563-573 (2002).
 [4]
 RN SEQUENCE FROM N.A. (ISOFORM 1).
 RP TISSUE=Breast;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Donald M.P., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalios D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 [5]
 RN SEQUENCE OF 671-869 FROM N.A.
 RP STRAIN=C57BL/6J; TISSUE=Forelimb;
 RX MEDLINE=20027348; PubMed=10556317;
 RA Wang A.J., Pierce A., Judson-Kremer K., Gaddis S., Aldaz C.M.,
 RA Johnson D.G., MacLeod M.C.;
 RT "Rapid analysis of gene expression (RAGE) facilitates universal
 expression profiling.";

Nucleic Acids Res. 27:4609-4618(1999).
 CC -!- FUNCTION: May play a role in the regulation of both apoptosis and
 CC cell proliferation.
 CC -!- SUBUNIT: Interacts with PDCD6/ALG-2 and SH3BP1 (By similarity).
 CC The interaction with PDCD6 is dependent on calcium.
 CC -!- ALTERNATIVE PRODUCTS;
 CC Event=Alternative splicing; Named isoforms=3;
 CC Name=1;
 CC IsoId=Q9WU78-1; Sequence=VSP 007501;
 CC Name=2; Synonyms=Alx-SF, Short;
 CC IsoId=Q9WU78-2; Sequence=VSP 007502;
 CC Note=Does not interact with ALG-2;
 CC Name=3;
 CC IsoId=Q9WU78-3; Sequence=VSP 007501;
 CC -!- TISSUE SPECIFICITY: Ubiquitously expressed.
 CC -!- SIMILARITY: Contains 1 BRO1 domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AJ005073; CAA06329.1; -
 CC EMBL; AJ005074; CAA06330.1; -
 CC EMBL; AF119955; AAD26813.1; -
 CC EMBL; AK031256; BAC27323.1; -
 CC EMBL; BC002261; AAH02261.1; -
 CC EMBL; BC026823; AAH26823.1; -
 CC EMBL; AF176514; AAD31115.1; -
 CC MGD; MGI:1333753; Pdc6ip.
 CC GO; GO:0005829; Cytoisol; IDA.
 CC InterPro; IPR004328; BRO1.
 CC Pfam; PF03097; BRO1; 1.
 CC Apoptosis; Alternative splicing.
 CC DOMAIN 1 167
 CC PRO-RICH.
 CC K -> KYFYFQ (in isoform 3).
 CC VARSPPLIC 239 239
 CC /FTid=VSP 007501.
 CC Missing (in isoform 2).
 CC /FTid=VSP 007502.
 CC CONFLICT 329 333
 CC LDPIG -> SGSYR (IN REF. 2).
 CC V -> L (IN REF. 2, 3 AND 4).
 CC CONFLICT 530 530
 CC EV -> DL (IN REF. 2).
 CC CONFLICT 547 548
 CC N -> T (IN REF. 4).
 CC CONFLICT 595 595
 CC L -> V (IN REF. 2).
 CC CONFLICT 625 625
 CC KQ -> NE (IN REF. 2).
 CC CONFLICT 640 641
 CC G -> R (IN REF. 2).
 CC CONFLICT 821 821
 CC P -> L (IN REF. 1; CAA06330).
 CC CONFLICT 853 853
 CC SEQUENCE 869 AA; 96010 MW; 9AA84B592FDCEASE CRC64;
 SQ
 Query Match 7.1%; Score 266.5; DB 1; Length 869;
 Best Local Similarity 26.1%; Pred. No. 5.9e-10;
 Matches 110; Conservative 60; Mismatches 151; Indels 101; Gaps 15;
 QY 167 LKTKELDWSTPLKELISVHF---GEGASVEAEIRRELEALRQAM--RTPSRNIEGLELL 221
 Db 9 LKTKSEVDLAKELVKFIQOTVPSGEEQACRAAEELSKLRSSALGPLDKHEGALFTL 68
 QY 222 TAYYNQLCFDLARFLTPARSLGLFFHWYD-----SLTG--VPAQORALAFKGSVLFNIG 274
 Db 69 LRYVDQICISIEPKFPFSENQICLTFTWKDAFDKSLFGSVKALASLGYEKSVCVFNCA 128
 QY 275 ALHTQIGARQDRCTEGARRAMEAFORAGAFSLIREN----FSHAPSPDMSAASLCALE 330
 Db 129 ALASQIAAEQNLDNDEGLKTKAAQYQFASGAFLEHKOTVLSALSREPTVDISPDVTGLS 188
 QY 331 OLHMAQAQCEVFEGLSPPASMAPQDCIAQLRLAQEAQAAYEYRLVHRTMAQPFVHDYVP 390
 Db 189 LIMLAQAQEVFF-----LKATRDKMKDAIIAKLANO-AADY--FGDAFKQCQYKOTLP 238

QY 391 VSWTALVHVKAEPFRSLAHYHVMALCDGSPATEGELPTEHQVFLQPTSSKRGVLPQ 450
Db 239 KEVPTLAAQCIQMAEAHYQSI-----262
QY 451 ELBERROLGKAHLKRAILGQEAHLHALCRVLRVLDLLRAV-----ISOTLQ 498
Db 263 -----LAKQOKX---FGBEIARLQHA-----ABLIKXNASRYDEYVNVKDFSKIN 305
QY 499 RSLAKYAELOREDDFCEAAEAPDIQ-----PKTHQKPEARMPLRSQKGPDIPIHRLGPL 552
Db 306 RAL-----TAARKNDNFIYHDRVLDLKDLPDKATLVKPTVNVPSQ-KFTDLFEKVPV 361
QY 553 SV 554
Db 362 SV 363

RESULT 6

ID_P061_XENLA STANDARD; PRT; 867 AA.
AC Q9W6C5;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DE Programmed cell death 6 interacting protein (Signal transduction protein Xp95).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
[1]
RN SEQUENCE FROM N.A. (ISOFORM 1), AND PHOSPHORYLATION.
RP MEDLINE=99150334; PubMed=10026166;
RA Che S., El-Hodiri H.M., Wu C.-F., Nelman-Gonzalez M., Weil M.M., Etkin L.D., Clark R.B., Kuang J.;
RT "Identification and cloning of Xp95, a putative signal transduction protein in Xenopus oocytes.";
RL J. Biol. Chem. 274:5522-5531(1999).
[2]
RN SEQUENCE FROM N.A. (ISOFORM 2).
RP TISSUE=Embryo;
RA Klein S., Strausberg R.L.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
CC !- FUNCTION: May be a signal transduction protein.
CC !- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q9W6C5-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q9W6C5-2; Sequence=VSP_007503;
CC !- PTM: Phosphorylated on tyrosine residues.
CC !- SIMILARITY: Contains 1 BR01 domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----

DR EMBL; AF115497; AAD20341.1; --
DR EMBL; BC043849; AAH43849.1; ALT_INIT.
DR InterPro: IPR004328; BR01.
DR Pfam; PF03097; BR01; 1.
KW Phosphorylation; Alternative splicing.
FT DOMAIN 1 166
FT PRO-RI. 166
FT DOMAIN 716 866
FT VARSPLIC 238 238
FT K -> KYFYFQ (in isoform 2).
FT /FTIG=VSP_007503.
SQ SEQUENCE 867 AA; 96198 MW; 6BE7173E68263B3 CRC64;

Query Match 6.9%; Score 261.5; DB 1; Length 867;
Best Local Similarity 24.5%; Pred. No. 1.3e-09;
Matches 106; Conservative 72; Mismatches 144; Indels 111; Gaps 16;
QY 158 VTPMPLGLGKEKELDWSPTLKLISVHF--GEDGASYEAEIRELEALQAM--RTPSR 213
Db 5 ISVP-----LKKTSEVDLVKPLSKYIHNTYPSGSDQTEYCRVDELNKLRSKSAVGRPLDK 59
QY 214 NESGLELLTAYYNOLCFDARFLTPARSLGLFPHWYDSLT-----GVPAQQBALAFEK 266
Db 60 HETSLETFWERYYOLCSVEPKFPFTESQLCTITFKDAFDKGSIFGGSVKLALPSLGVK 119
QY 267 GSVLFNIGALHTQIGARQDRSCTEGARRAMEAFORAGAFSLIRE----NFSHAPSMDMS 322
Db 120 TCVLFNIGALASQIASQNLDNDEALKAAKFKYQLASGAFSHIKDTVTLSSLNRPDTVDIS 179
QY 323 AASLCALAEOLMMAQAECEVFEGLSPSPASMAPODCLAQLRLAQAAQVAAEYRLVHRTMAQ 382
Db 180 PDIVGTLTLINLAQAQEVFF-----LKATRDMDKDAVIKLANQA-----220
QY 383 PPVHDYVPSVSWTALVHVKAEPFRSLAHYHVMALCDGSPATEGELPTEHQVFLQPTSSK 442
Db 221 ----DY-----YGDFAKQCQ-----YKDTLSK 238
QY 443 PRGVLPQLEERRQLGKAHLK--RAILGQEAHLHALCRVLRVLDLLRAV-----492
Db 239 EVFPI-----LAAXHCIMQAAHAEVHQSVLAKQOKFGEEIGRLQHASLDLVKTVSSRYDEYV 294
QY 493 ----ISOTLQSLAKYAELOREDDFCEAAEAPDI-----QPKTHQKPEARMPLRSQK 541
Db 295 NVKDLADKINRAL---TAARKNDNFIYHDRVLDLKDLPDKATLVKPTVNVPSV--LSQ-K 349
QY 542 GPDIFHRLGPLSV 554
Db 350 YTDLFEKVPVLA 362

RESULT 7

BR01 YEAST STANDARD; PRT; 844 AA.
ID BR01 YEAST
AC P48582; 002823;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE BR01 protein.
GN BR01 OR YPL084W OR LPF2.
OS Saccharomyces cerevisiae (Baker's Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96220442; PubMed=8649366;
RA Nickas M.E., Yaffe M.P.;
RT "BR01, a novel gene that interacts with components of the Pkc1p-mitogen-activated protein kinase pathway in Saccharomyces cerevisiae.";
RL Mol. Cell. Biol. 16:2585-2593(1996).
[2]
RN SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RX MEDLINE=973113271; PubMed=9169875;
RA Bussey H., Storms R.K., Ahmed A., Albermann K., Allen E., Ansong W., Araujo R., Aparicio A., Barrell B.G., Badcock K., Benes V., Botstein D., Bowman S., Bruckner M., Carpenter J., Cherry J.M., Chung E., Churcher C.M., Coster F., Davis K., Davis R.W., Dietrich F.S., Delius H., DiPaolo T., Dubois E., Duesterhoeft A., Duncan M., Floeth U., Fortin N., Friesen J.D., Fritz C., Goffeau A., Hall J., Hebling U., Heumann K., Hilbert H., Hillier L.W., Hunnicke-Smith S., Hyman R., Johnston M., Kalman S., Klein K., Komp C., Kurdi O., Lashkari D., Lew H., Lin D., Louis E.J., Marathe R., Messenguy F., Mewes H.-W., Mirtipati S., Moestl D., Mueller-Auer S., Namath A., Nentwich U., Oefner P., Pearson D.,

NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
MEDLINE=95080426; PubMed=7988719;
Palmer R.H., Ridden J., Parker P.J.;
"Identification of multiple, novel, protein kinase C-related gene
products.";
FEBS Lett. 356:5-8(1994).
[2]
SEQUENCE FROM N.A.
TISSUE=Retal brain;
MEDLINE=95154310; PubMed=7851406;
Palmer R.H., Ridden J., Parker P.J.;
"Cloning and expression patterns of two members of a novel protein-
kinase-C-related kinase family.";
Eur. J. Biochem. 227:344-351(1995).
[3]
SEQUENCE FROM N.A., AND MUTAGENESIS OF LYS-644.
TISSUE=Hippocampus;
MEDLINE=94183274; PubMed=8135837;
Mukai H., Ono Y.;
"A novel protein kinase with leucine zipper-like sequences: its
catalytic domain is highly homologous to that of protein kinase C.";
Biochem. Biophys. Res. Commun. 199:897-904(1994).
-1- FUNCTION: CAN PHOSPHORYLATE RIBOSOMAL PROTEIN S6. MEDIATES GTPASE
RHO DEPENDENT INTRACELLULAR SIGNALING (BY SIMILARITY).
-1- ENZYME REGULATION: Activated by lipids, particularly cardiolipin
and to a lesser extent by other acidic phospholipids (By
similarity).
-1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
-1- TISSUE SPECIFICITY: Found ubiquitously. Expressed in heart, brain,
placenta, lung, skeletal muscle, kidney and pancreas.
-1- PTM: Autophosphorylated; preferably in serine.
-1- PTM: Activated by limited proteolysis with trypsin (By
similarity).
-1- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. PKC
subfamily.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; U33053; AAC50209.1; -.
EMBL; S75546; AAB33345.1; -.
EMBL; D26181; BAA05169.1; -.
PIR; JCI2129; JCI2129.
PDB; 1CXZ; 23-MAY-00.
Genew; HGNC:9405; PRKCL1.
MIM; 601032; -.
GO; GO:0004672; F:protein kinase activity; TAS.
GO; GO:0007257; P:activation of JUNK; TAS.
GO; GO:0006468; P:protein amino acid phosphorylation; TAS.
GO; GO:0007165; P:signal transduction; TAS.
InterPro; IPR000008; C2.
InterPro; IPR008973; C2_CaLB.
InterPro; IPR000961; Pkinase.C.
InterPro; IPR000719; Prot_kinase.
InterPro; IPR000861; REM_repeat.
InterPro; IPR008271; Ser_thr_pkin_AS.
InterPro; IPR002290; Ser_thr_pkinase.
Pfam; PF02185; HRI; 3.
Pfam; PF00069; pkinase; 1.
Pfam; PF00433; pkinase_C; 1.
ProDom; PD000001; Prot_kinase; 1.
SMART; SM00239; C2; 1.
SMART; SM00074; HRI; 3.
SMART; SM00133; S_TK_X; 1.
SMART; SM00220; S_TKC; 1.
PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.

DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
 DR PROSITE; PS00111; PROTEIN KINASE DOM; 1.
 KW Transferase; ATP-binding; Serine/threonine-protein kinase;
 KW Phosphorylation; Polymorphism; 3D-structure.
 FT DOMAIN 615 874 PROTEIN KINASE.
 FT NP BIND 621 629 ATP (BY SIMILARITY).
 FT BINDING 644 644 ATP (BY SIMILARITY).
 FT ACT SITE 740 740 BY SIMILARITY.
 FT VARIANT 901 901 V->I (in dbSNP:10846).
 FT /FTID=VAR_014937.
 FT MUTAGEN 644 K-R: SUBSTANTIAL REDUCTION OF
 FT AUTOPHOSPHORYLATION.
 FT CONFLICT 191 191 D->G (IN REF. 3).
 SQ SEQUENCE 942 AA; 103989 MW; A89E40DCAEF560E3 CRC64;
 Query Match 4.1%; Score 156.5; DB 1; Length 942;
 Best Local Similarity 21.4%; Pred. No. 0.012;
 Matches 159; Conservative 79; Mismatches 257; Indels 247; Gaps 37;
 69 QLHAVGCDSLTQIQGQLQRRAGIHOQIDKELQMRGAEINLYRATSN-NRVRETVALEL 127
 18 QLGLAGADLAAPGVQQQLERERLRREIRKELKKEGAENLRATTDLGRSLGPEVLL 77
 128 SYVNSNQLLKELEELSGGV---DP-GRHGEAVTVPMPLGLKELKELDMSTPLKELI 183
 78 RGSRRDLQLHQQLQELHAAHVLPDPAATHDG-----PQSP----- 113
 184 SVHGEQASVEA-EIREALRQAMTPSPNESGLELLTAYYNQLCFDLARFLTPARSL 242
 114 ----GAGGPTCSATNLSRVAGLEKQALAEIKVKQGAENMIQTNGSTKDRKLLTAQOM 169
 243 GLFPHWYDSLTVGP-----AQRALAFE-----KGVLFNIGALHTQIGA----- 282
 170 ----LQDSKTKIDILRMQLRALQADOLENQAAAPDQTQSP--DLGAVELRIEELRHHF 222
 283 RQDRSCTEGARAMEAFQRAAGAFSLIRENFSPHAPSMDMSAASICALBQLMMAQAQBCVF 342
 223 RVEHAVAEGAKNVLRL-----LSAAKAPDRKAVSEAAQEKLKTESNOKLGLLR 268
 343 EGLSPASMAPQDCIAQLRLAQEAQVAAYELVHRTWAQPPVHDY----VPVSWTALVH 398
 269 EALERLIGELPAD-HPKGLRLRELAASAASAFSTRLAGPPFATHYSTLCKPAPLTGTL 327
 399 VKAEYFRSLAH---YHVAMALCDGSPATEGELPTHEQVFLQPP----- 438
 328 VRVVGCRDLPETIPWNPFSM--GGPGI-----PDSRPFLSRPARGLYSRSGSLGRSSL 381
 439 -----TSSKPRGP-----VLPQLEERRQLGKAHLKRAILGQ 470
 382 KAEANTSEVSTVLKLDNTVVGQTSWKPCGPNAMDQSFTELEARE-----LELAVFWR 436
 471 EEARLHALC--RVLRVDLRAVISQTLQRLSLAKYAELEDREDDFCEAAEAPDIQPKTHQ 528
 437 DO-----RGLCALFKLKEDEL-----DNERHEVQDMPEQCLVAEV-----TFR 477
 529 KPE-ARMPRL-----SQCKGPDIFHRLGFLSVFSAKRWRLVGPVHLTR----- 571
 478 NPVIERIPLRRQKKIFSKQQGA---FQRAEQNMIDVA--TW-----VRLRLRLIPNAT 527
 572 GEGGFLTLRGDSPVLIIAVTPGSAAGLKEGDIYIVSNGQPCRWWRHAEVVTLEKAA 631
 528 GTGTF-----SP-----GASFGSARTTG-----DISVE-----KINLG 556
 632 GEAGASLOWVSLPLSSRLPSLGRDPRVLLGPRGLLRSQREHGCKTPASTWASPRLLNWS 691
 557 TDSDESPOKSRDPPSPSSLS--PI-----QESTAPELPSFTQETPGAL--- 601
 632 RKAQQKTKGGCQPCAPVKPAP 713
 602 -----CSPLRKSP 609

PKL1 RAT
 ID PKL1 RAT STANDARD; PRT; 946 AA.
 AC Q63433;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Protein kinase C-like 1 (BC 2.7.1.-) (Protein-kinase C-related kinase
 DE 1) (Protein kinase C-like PKN) (Serine-threonine protein kinase N)
 DE (Protease-activated kinase 1) (PAK-1).
 GN P3KCL1 OR P3K1 OR PKN.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lung;
 RX MEDLINE=94183274; PubMed=8135837;
 RA Ono Y., Mukai H.;
 RT "A novel protein kinase with leucine zipper-like sequences: its
 RT catalytic domain is highly homologous to that of protein kinase C.";
 RL Biochem. Biophys. Res. Commun. 199:897-904(1994).
 RN [2]
 RP CHARACTERIZATION.
 RC TISSUE=Liver;
 RX MEDLINE=94327556; PubMed=8051089;
 RA Morrice N.A., Gabrielli B., Kemp B.E., Wettenhall R.E.;
 RT "A cardiolipin-activated protein kinase from rat liver structurally
 RT distinct from the protein kinases C.";
 RL J. Biol. Chem. 269:20040-20046(1994).
 CC -!- FUNCTION: CAN PHOSPHORYLATE RIBOSOMAL PROTEIN S6. MEDIATES GTPASE
 CC RHO DEPENDENT INTRACELLULAR SIGNALING.
 CC -!- ENZYME REGULATION: Activated by lipids, particularly cardiolipin
 CC and to a lesser extent by other acidic phospholipids.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -!- PM: Autophosphorylated; preferably in serine.
 CC -!- PTM: Activated by limited proteolysis with trypsin.
 CC -!- SIMILARITY: Belongs to the Ser/thr family of protein kinases. PKC
 CC subfamily.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
 CC or send an email to license@isb-sib.ch).

 CC EMBL; D26180; BAA05168.1; -;
 CC HSSP; P05132; 1ATP.
 CC InterPro; IPR000008; C2.
 CC InterPro; IPR008973; C2 CalB.
 CC InterPro; IPR000961; Pkinase_C.
 CC InterPro; IPR000719; Prot_kinase.
 CC InterPro; IPR000861; REM_repeat.
 CC InterPro; IPR008271; Ser_thr_pkin_AS.
 CC InterPro; IPR002290; Ser_thr_pkinase.
 CC Pfam; PF02185; HR1; 3.
 CC Pfam; PF00069; pkinase; 1.
 CC Pfam; PF00433; pkinase_C; 1.
 CC ProDom; P000001; Prot_kinase; 1.
 CC SMART; SM00239; C2; 1.
 CC SMART; SM00074; HR1; 3.
 CC SMART; SM00133; S_TK_X; 1.
 CC SMART; SM00220; S_TK_X; 1.
 CC PROSITE; PS00108; PROTEIN KINASE ST; 1.
 CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 CC PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
 CC Transpherase; ATP-binding; Serine/threonine-protein kinase;
 KW Phosphorylation.
 FT DOMAIN 615 874 PROTEIN KINASE.
 FT NP BIND 625 633 ATP (BY SIMILARITY).
 FT BINDING 648 648 ATP (BY SIMILARITY).


```
FT ACT SITE 744 744 BY SIMILARITY.
SQ SEQUENCE 946 AA; 104468 MW; FE9D4EFP8E9EBA4B CRC64;

Query Match
Best local similarity 3.9%; Score 146.5; DB 1; Length 946;
Matches 154; Conservative 21.2%; Pred. No. 0.058;
Matches 154; Conservative 271; Indels 211; Gaps 34;

QY 69 QLVAVGDSLTQICGQLOSRRAHQIQIDKELQMRGAENLYPATSN-NRVRETVALEL 127
DB 18 QLVAGADLAAPGVQQQLERERELREKELKEGAENLRATTDLGRSLAPVELL 77
QY 128 SYVNSNLQLLKEELELELSGGV---DPRGHGEAVTVPMIPLGKETKELDWSLTKELIS 184
DB 78 RGSARRLDLHQLOELHAHVLPDP-TAGSDA-----FQSLAEGSPVCSNLSRVAG 130
QY 185 VHGEGDASVEAREIREALRQAMRTSPRNSGLELITAYYNQCLFLDRLFTPARSLGL 244
DB 131 L--EKQLATELVKQ-----GAENMIQYNSGSTKDKLLLTAAQOM-- 169
QY 245 FFHWYSLTGVV---AQORAL-AFEKGSVLF-----NIGALHTQIGA---R 283
DB 170 ---LQSSKTKIDIRMQLRALQALQAGLESQAAPDEAHGDPDLGAVELRIEELRHFR 226
QY 284 QDRSCTEGAR-----RAMEAFQRAAGAFSLIRENFESH-----APSP 319
DB 227 VEHAVAEAGAKNVLRLLSAAKAPDRKAVSEAQEKLTESNQKLGILRESLRLGELPADHP 286
QY 320 -----DMSAASLCALQLMMAQAECEVPEGISPPASMAPQDCLAQLRLAQBAAQVAEE 372
DB 287 KGLRLREELTAASAAFSAILPGFPFATHYSTLSKPAFLT-----GTLEVRVVG 336
QY 373 YRLVHRTMAQPEVH---DYVPVSWTALVHVKAEB--YFRSLAHYHVAMALCDGSPATEGEL 427
DB 337 KNLPEIIPSPSPSVGASGTPDRTPLSRPARGLYNRS-----GSLSGRSSLKGEA 388
QY 428 PTHEQV-----FLQPTTSXPRGP-----VLPQELERRQLGKAHLKRAILGQEEAL 474
DB 389 ENSTVSTVLKLDNTVVQGTAWKPCGNFNAWDQSTLELEARE-----LELAVFWRDQ-- 441
QY 475 RLHALC--RVLRVDLLRAVISOTLORSIAKYAELOREDDFCSAABAPDIQPKTHQK-P-E 531
DB 442 --RGLCALFKLEDFL-----DNERHEVQLDMEPQGLVAEV-----TFNPPII 484
QY 532 ARMPRLSQGKGPDIPIRLGLPLSVFSAKNRWLVGPVHLTRGEGGFGILTRGDSFVLIAAV 591
DB 485 ERIPRLQROK-----KIFSQQGQTFORARHMINIVATWVRLLR-----RL 525
QY 592 IFGSQAAAAGLKEGDYIVSVNGQPCFWRHRAEVVTVTELKAAGEAGASIQVLSLPPSRPLPS 651
DB 526 IPNAVATGS-----FSPNASP-----GSEIRSTGD--ISWEKLNLGADS--DS 564
QY 652 LGDRRPVLGPRGLLRQREHGCKT-----PASTWASPRLLNWSRKAQQGKTGGCPQPCA 707
DB 565 SQSKGSPAGLPSTSCSLSPSTSTTSPELSETOETPGPGL-----CS 607
QY 708 PVKPAP 713
DB 608 PLRKSP 613

RESULT 10
EMIL MOUSE
ID EMIL MOUSE STANDARD; PRT: 1017 AA.
AC Q99K41;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE EMILIN 1 precursor (Elastin microfibril interface-located protein 1)
DE (Elastin microfibril interfacer 1).
GN EMILIN1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
```

CC nucleation of the triple helix and then a further quaternary
 CC assembly to higher order polymers via intermolecular disulfide
 CC bonds. Interacts with EMILIN2 (By similarity).
 CC -!- SUBCELLULAR LOCATION: Extracellular matrix. Found mainly at the
 CC interface between amorphous elastin and microfibrils.
 CC -!- DEVELOPMENTAL STAGE: Detectable in morula and blastocyst. First
 CC expressed in ectoplacental cone in embryos of 6.5 days and in
 CC extraembryonic visceral endoderm at 7.5 days. Expressed also in
 CC the allantois. Expression in the ectoplacental cone-derived
 CC secondary trophoblast giant cells and spongiotrophoblast is strong
 CC up to 11.5 days and then declines. In the embryo, high levels are
 CC initially expressed in blood vessels, perineural mesenchyme and
 CC somites at 8.5 days. Later on, intense expression is identified in
 CC the mesenchymal component of organs anlage (ie lung and liver) and
 CC different mesenchymal condensations (ie limb bud and branchial
 CC arches). At late gestation expression is widely distributed in
 CC interstitial connective tissue and smooth muscle cell-rich
 CC tissues.
 CC -!- SIMILARITY: Contains 1 C1Q domain.
 CC -!- SIMILARITY: Contains 1 collagenous domain.
 CC -!- SIMILARITY: Contains 1 EMI domain.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; AK029337; AAC26403.1; -;
 DR EMBL; BC005481; AAH05481.1; -;
 DR MGD; MGI:1926189; Emilin1.
 DR InterPro; IPR001073; C1q.
 DR InterPro; IPR008160; Collagen.
 DR Pfam; PF00386; C1q; 1.
 DR Pfam; PF01391; Collagen; 1.
 DR Cell adhesion; Extracellular matrix; Signal; Coiled coil; Collagen;
 KW Glycoprotein. 1 23
 FT SIGNAL 24 1017 POTENTIAL.
 FT CHAIN 57 131 EMILIN 1.
 FT DOMAIN 171 211 COILED COIL (POTENTIAL).
 FT DOMAIN 237 266 COILED COIL (POTENTIAL).
 FT DOMAIN 269 275 POLY-GLY.
 FT DOMAIN 310 374 COILED COIL (POTENTIAL).
 FT DOMAIN 519 573 COILED COIL (POTENTIAL).
 FT DOMAIN 676 697 COILED COIL (POTENTIAL).
 FT DOMAIN 789 809 COILED COIL (POTENTIAL).
 FT DOMAIN 815 865 COLLAGEN.
 FT DOMAIN 873 1008 C1Q.
 FT CARBOHYD 156 156 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 416 416 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 456 456 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 562 562 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 659 659 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 767 767 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 795 795 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT SEQUENCE 1017 AA; 107584 MW; 2EF903204DB9C88F CRC64;

Query Match
 Best Local Similarity 21.0%; Score 146; DB 1; Length 1017;
 Matches 183; Conservative 87; Mismatches 277; Indels 324; Gaps 45;
 47 WRGCRVHGED-----PVRVHGVPMNPQLHANGCDSLTQICCG 85
 120 WRCCGYGDCGEGPASVLPAPSTPLPRPRPRPNLSGSSAGSLGLGGGPGVESEK 179
 86 LQSRRAQIHQIDKELQMTG-----AENLYRATS-----VRETVALE 126
 180 VQQLERQV-KSLTKELQGLGVLMNGRLAEDVQRAVDTVFNGRQOPADAAARPGVHET 238

QY 127 LSYVNSLQLL-----KEELEL-----SGVDPGRHSGEAVTVPMIPLGKTKETKELDW 175
 Db 239 LSEIQOQLLDNRVSTHDQELGLNHNHNG--PGGGGRASGPVP-VPSGPFSEELLRL 295
 QY 176 STPLKELLSVHF-CEDGASYAE-----IRELEALQAM----- 208
 Db 296 ERQLQSCSVCLTGLDGFPROQQQEDREKLRTLEKLMSSMEERQOOLVCPAMARPPQSCC 355
 QY 209 -----RTPSRNESGLELL-----TAYYNQLCFELDARF 235
 Db 356 PPELGRVSELERLDVVTGSLTVLSGRGSELGGAAGGHPGYTSLASRLSLEDRLF 415
 QY 236 ---LTPARSL-----GLFFHWYDLSLTGPAQORALAFKGVSLFNGALHTGIGARQ 284
 Db 416 NSTLGPSEBEQKNWPGGRLGHM---LPAAPGRLEKLE---GLLANVS---RELGGRM 465
 QY 285 DRSCTEGARAMEAFQRAAGFSLRLRENFSHAPSPDMSAAS--LCALQLQMM----- 334
 Db 466 D-----LLE--EQVAGAVRTCGQICSGAPEQDSRVNEILSALERRVLDSEGRQL 514
 QY 335 -----AOAQECVFEGSLPPSPASMAPQDCLAQLAQLAQA-AQVAAEYVL-VHRTVA 381
 Db 515 VSGGLHEAAEAGEAQAVLEGL-----QGLLSRLRERMDAQEBTAAEILLRLNLTA 566
 QY 382 QPPVHDYVFWNTALVHVKAIFYRSLAHYHVMALCDGSPATEGEL-----PTHEQ 432
 Db 567 QLS-----OLEGLLQARGD-----EGCGACGVQBELGRLRDGVERCSCP--- 606
 QY 433 VFLOPP-----TSSKPRGPVLPQLEERRQLGKAHLKRAILQGEALRLHALCRVLR 484
 Db 607 --LLPPRGAGPGVGGFSRGL-----DGFSVFGSGSSALQALQGLS 649
 QY 485 EVDLLRAVISQTLQSLAKYAELEDREDDFCEAAEAPDIOPTHQKPEARMPLRSOGKGP 544
 Db 650 EVILTFSSLSLH-----ELQTTVEGQAGD 675
 QY 545 IPHRLGPL--SVFSAKNRWLVGPVHLTRGEGGF-----GLTLRGDSPLV----- 587
 Db 676 LAD-LGATKDSIISINLQOEATHEVTESEERFRGLEBEGQAQGCPSLEGRLRLEGV 734
 QY 588 ---IAAVIPGQAAAAGLKEGDYIVSVNGQPCRMW---RHAENVTELKAA-----GEA 634
 Db 735 CERLDTVAGLQ---GLREG-----LSRHVAGLAAVRESNSTSLTQAALLKLLGGQA 785
 QY 635 GASLQVSVLLPSSRLPSLGDRPVL-----LQPRGLLSQRHGHCKTPTASTWASPLPLN 689
 Db 786 GLGRRLLGAL--NNSLLLEDLRLQSLKDFTPSGKAGPPGPPGLQGPSGAPGPPGPP-- 841
 QY 690 WSRKAQOGKTGCPQP-----CAPVKEAP 713
 Db 842 -GKDGQQAIGP-PGPGQGAEGAPAPVP 870

RESULT 11

ARBH HUMAN STANDARD; PRT; 1522 AA.
 AC O15085;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Rho guanine nucleotide exchange factor 11 (PDZ-RhoGEF).
 GN ARHGEF11 OR KIAA0380.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=97349984; PubMed=9205841;
 RA Nagase T., Ishikawa K.-I., Nakajima D., Chira M., Seki N.,
 RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. VII."

RT The complete sequences of 100 new cDNA clones from brain which can
 RL code for large proteins in vitro.";
 RN DNA Res. 4:141-150(1997).

[2]

RP INTERACTION WITH RHOA.

RX MEDLINE=99456836; PubMed=10526156;

RA Ruemenapp U., Blomquist A., Schwoerer G., Schablowski H., Psoma A.,

RA Jakobs K.H.;

RT "Rho-specific binding and guanine nucleotide exchange catalysis by
 RT KIAA0380, a dbl family member.";

RL FEBS Lett. 459:313-318(1999).

[3]

RN INTERACTION WITH GNA12 AND GNA13, AND TISSUE SPECIFICITY.

RP

RC TISSUE=Brain;

RX MEDLINE=99150378; PubMed=10026210;

RA Fukuhara S., Murga C., Zohar M., Igishi T., Gutkind J.S.;

RT "A novel PDZ domain containing guanine nucleotide exchange factor

RT links heterotrimeric G proteins to Rho.";

RL J. Biol. Chem. 274:5868-5879(1999).

[4]

RN INTERACTION WITH PLXNB1 AND PLXNB2.

RX MEDLINE=2225992; PubMed=12372594;

RA Driessens M.H., Olivo C., Nagata K., Inagaki M., Collard J.G.;

RT "B plexins activate Rho through PDZ-RhoGEF.";

RL FEBS Lett. 529:168-172(2002).

[5]

RN INTERACTION WITH PLXNB1 AND PLXNB2.

RX MEDLINE=22301907; PubMed=12183458;

RA Ferrot V., Vazquez-Prado J., Gutkind J.S.;

RT "Plexin B regulates Rho through the guanine nucleotide exchange

RT factors leukemia-associated Rho GEF (LARG) and PDZ-RhoGEF.";

RL J. Biol. Chem. 277:43115-43120(2002).

[6]

RN SUBCELLULAR LOCATION.

RX MEDLINE=20449035; PubMed=10900204;

RA Togashi H., Nagata K., Takagishi M., Saitoh N., Inagaki M.;

RT "Functions of a rho-specific guanine nucleotide exchange factor in
 RT neurite retraction. Possible role of a proline-rich motif of KIAA0380
 RT in localization.";

RL J. Biol. Chem. 275:29570-29578(2000).

[7]

RN X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 281-490.

RX MEDLINE=21364065; PubMed=11470431;

RA Longenecker K.L., Lewis M.E., Chikumi H., Gutkind J.S.,

RA Derewenda Z.S.,

RT "Structure of the RGS-like domain from PDZ-RhoGEF: linking

RT heterotrimeric g protein-coupled signaling to Rho GTPases.";

RL Structure 9:559-569(2001).

CC -!- FUNCTION: May play a role in the regulation of RhoA GTPase by

CC guanine nucleotide-binding alpha-12 (GNA12) and alpha-13 (GNA13).

CC Acts as guanine nucleotide exchange factor (GEF) for RhoA GTPase

CC and may act as GTPase activating protein (GAP) for GNA12 and

CC GNA13.

CC -!- SUBUNIT: Interacts with GNA12 and GNA13 through the RGS domain.

CC Interacts with RHOA, PLXNB1 AND PLXNB2. Interacts with SLC1A6 (By

CC similarity).

CC -!- SUBCELLULAR LOCATION: Cytoplasmic. Translocated to the membrane

CC upon stimulation.

CC -!- TISSUE SPECIFICITY: Ubiquitously expressed.

CC -!- DOMAIN: The poly-Pro region is essential for plasma membrane

CC localization upon stimulation.

CC -!- SIMILARITY: Contains 1 DBL-homology (DH) domain.

CC -!- SIMILARITY: Contains 1 PDZ/DHR domain.

CC -!- SIMILARITY: Contains 1 PH domain.

CC -!- SIMILARITY: Contains 1 RGS (RGS-like) domain.

CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; AB002378; BAA20834.1; -.
 DR PDB; 1HTJ; 11-JUL-01.
 DR Genew; HGNC:14580; ARHGEF11.

DR MIM; 605708; -.

DR InterPro; IPR001331; GDS_CDC24.

DR InterPro; IPR001478; PDZ.

DR InterPro; IPR001849; PH.

DR InterPro; IPR000342; Regl Gproteins.

DR InterPro; IPR000219; RhoGEF.

DR Pfam; PF00595; PDZ; 1.

DR Pfam; PF00621; RhoGEF; 1.

DR SMART; SM00228; PDZ; 1.

DR SMART; SM00233; PH; 1.

DR SMART; SM00315; RGS; 1.

DR SMART; SM00325; RhoGEF; 1.

DR PROSITE; PS00741; DH 1; FALSE_NEG.

DR PROSITE; PS50010; DH 2; 1.

DR PROSITE; PS50106; PDZ; 1.

DR PROSITE; PS50003; PH_DOMAIN; 1.

DR PROSITE; PS50132; RGS; 1.

KW Guanine-nucleotide releasing factor; GTPase activation; Coiled coil;

KW 3D-structure.

FT DOMAIN 47 126 PDZ.

FT DOMAIN 306 486 RGS.

FT DOMAIN 444 470 COILED COIL (POTENTIAL).

FT DOMAIN 734 923 DH.

FT DOMAIN 965 1079 PH.

FT DOMAIN 1094 1099 POLY-PRO.

SQ SEQUENCE 1522 AA; 167703 MW; CA16L25B9F8A4AA CRC64;

Query Match 3.9%; Score 145.5; DB 1; Length 1522;

Best Local Similarity 31.9%; Pred.No. 0.13;

Matches 51; Conservative 16; Mismatches 56; Indels 37; Gaps 6;

QY 567 VHLTRGEGGGLTLRGDSVLTAANVPGSQAAGLKEGDIYVNVGQPCRRWRHAEVVT 626

Db 48 VTIQDHQHGFGFTVSGDRIVLVQSVRPGGAAMKAGVKEGDIKIKVNGTMTNSSHLEVK 107

QY 627 ELKAGEAGASLQVWSLLPSSRLPSLGRRLPVLGRLRSQREHGCKTPTASTWASRP 686

Db 108 LIK-----SGAYV-ALTLLGSS-----PSSMGISGLQDPSAGAPRITSVIPSPPP 153

QY 687 LNWSRKAQGGTGGCPQCAPVK-----PAP---PSSLKH 719

Db 154 -----PPPLPPPPQRTGPKPLQDPEVQKH 177

RESULT 12

US1C_MOUSE

ID US1C_MOUSE STANDARD; PRT; 910 AA.

AC Q9ES64; Q91XD1; Q9CVG7; Q9ES65;

DT 10-OCT-2003 (Rel. 42, Created)

DT 10-OCT-2003 (Rel. 42, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Harmonin (Usher syndrome 1C protein homolog) (PDZ domain-containing

DE protein).

GN USH1C.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxId=10090;

RN [1]

RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 3), TISSUE SPECIFICITY, AND

RP ALTERNATIVE SPLICING.

RC TISSUE=Inner ear;

RX MEDLINE=20428180; PubMed=10973247;

RA Verpy E., Leibovici M., Zwaenepoel I., Liu X.-Z., Gal A., Salem N.,

RA Mansour A., Blanchard S., Kobayashi I., Keats B.J.B., Slim R.,

RA Petit C.;

RT "A defect in harmonin, a PDZ domain-containing protein expressed in

RT the inner ear sensory hair cells, underlies Usher syndrome type 1C.";

RT Nat. Genet. 26:51-55(2000).

```

[2]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RA Verpy E., Leibovici M., Zwaenepoel I., Blanchard S., Petit C.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RN TISSUE=Colon;
RC MEDLINE=22386257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Schenck C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schenck C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullihy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grinwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalios D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[4]
RP SEQUENCE OF 740-852 FROM N.A. (ISOFORM 1).
RC STRAIN=C57BL/6J; TISSUE=Small intestine;
RX MEDLINE=22354683; PubMed=12466851;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusci V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grilmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasaki Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Nunata K., Okido T., Pavan W.J., Perle G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sanderlin A., Schneider C., Sempile C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahlstedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hironaka-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aikawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imorani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
[5]
RP SUBUNIT: Associates with USH1G/SANS (By similarity).
RC -!- ALTERNATIVE PRODUCTS;
RA Event=Alternative splicing; Named isoforms=3;
RA Comment=Additional isoforms exist;
CC Name=3; Synonyms=b3;
CC IsoId=Q9ES64-1; Sequence=Displayed;
CC Name=2; Synonyms=b2;
CC IsoId=Q9ES64-2; Sequence=VSP_050532, VSP_050534;
CC Name=1; Synonyms=a1;
CC IsoId=Q9ES64-3; Sequence=VSP_050530, VSP_050531, VSP_050533;
CC -!- TISSUE SPECIFICITY: Isoform 1 is expressed in the eye, cochlea,
CC vestibule, heart, kidney, small intestine and testis; it is barely
```

Db 766 LALEGGVDSFVGKVVSAVYEGGAERHGGVKGDEIWAING-----KIVDTYTLA 816
QY 632 GEAGASLQ-----VSLPLS-----SRPLSGDRRPVL----- 659
Db 817 -EAAALQKAWNOGDWIDLWAVCPKPEYDDDELSSLPSSAABSPQLAKQLEAYEPVCR 875
QY 660 -----LGPRLRLRSQREHGCKTPASTWASPRP 686
Db 876 HGFFLQLEPTLLKSKERNQTFDSWRPASPAP 908

RESULT 13
ARHB_RAT
ID ARHB_RAT STANDARD; PRT; 1527 AA.
AC Q9ES67;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Rho guanine nucleotide exchange factor 11 (RhoGEF glutamate transport modulator GTRAP48).
GN ARHGEF11.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=101116;
RN [1]
SEQUENCE FROM N.A., FUNCTION, AND INTERACTION WITH RHOA; GNA13 AND SLC1A6.
RX MEDLINE=21133160; PubMed=11242047;
RA Jackson M., Song W., Liu M.-Y., Jin L., Dykes-Hoberg M., Lin C.-L.G., Bowers W.J., Federoff H.J., Sternweis P.C., Rothstein J.D.;
RT "Modulation of the neuronal glutamate transporter EAAT4 by two interacting proteins";
RL Nature 410:89-93(2001).
CC -!- FUNCTION: May play a role in the regulation of Rho GTPase by guanine nucleotide-binding alpha-12 (GNA12) and alpha-13 (GNA13). Acts as guanine nucleotide exchange factor (GEF) for RhoA GTPase and may act as GTPase activating protein (GAP) for GNA12 and GNA13 (By similarity).
CC -!- SUBUNIT: Interacts with RHOA, GNA13 and SLC1A6. Interacts with GNA12, PLXNB1 and PLXNB2 (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic. Translocated to the membrane upon stimulation (By similarity).
CC -!- SIMILARITY: Contains 1 DBL-homology (DH) domain.
CC -!- SIMILARITY: Contains 1 PDZ/DHR domain.
CC -!- SIMILARITY: Contains 1 PH domain.
CC -!- SIMILARITY: Contains 1 RGS (RGS-like) domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF225961; AAG28597.1; --
DR HSP; Q12923; 3PDZ.
DR InterPro; IPR001331; GDS_CDC24.
DR InterPro; IPR001478; PDZ.
DR InterPro; IPR001849; PH.
DR InterPro; IPR000342; Regl_Gprotein.
DR InterPro; IPR000219; RhoGEF.
DR Pfam; PF00595; PDZ; 1.
DR SMART; SM00621; RhoGEF; 1.
DR SMART; SM00228; PDZ; 1.
DR SMART; SM00233; PH; 1.
DR SMART; SM00315; RGS; 1.
DR SMART; SM00325; RhoGEF; 1.
DR PROSITE; PS00741; DH_1; FALSE_NEG.
DR PROSITE; PS50010; DH_2; 1.
DR PROSITE; PS50106; PDZ; 1.
DR PROSITE; PS50003; PH_DOMAIN; 1.

DR PROSITE; PS50132; RGS; 1.
KW Guanine-nucleotide releasing factor; GTPase activation; Coiled coil.
FT DOMAIN 64 143 PDZ.
FT DOMAIN 323 503 RGS.
FT DOMAIN 461 487 COILED COIL (POTENTIAL).
FT DOMAIN 742 931 DH.
FT DOMAIN 973 1087 PH.
SQ SEQUENCE 1527 AA; 168533 MW; ABAEA20F541A3A9A CRC64;
Query Match 3.6%; Score 136.5; DB 1; Length 1527;
Best Local Similarity 29.2%; Pred. No. 0.51; Mismatches 35; Gaps 4;
Matches 45; Conservative 18;
QY 567 VHLTRGEGGGLRLGDSPLVIAAVIPGSOAAAGLKEGDYIVSVNGQPCRMWRHAEVVT 626
Db 65 VLIQKDHGFGFTVSGDRIVLVQSVRPGGAAMKAGVKEGDRIIKVGNGTWTNSSHLEVVK 124
QY 627 ELKAAEGAGASQVSVLLPSSRLPSLGDRPVLLGPRLLRSQREHGCKTPASTWASPRP 686
Db 125 LKSGAYA-----ALTLLGSS-----PPSVGVSGLQNPDSVAGV-----LRVNP 163
QY 687 LLNWSRKAQGGTGGCPQCAPVKPAPPSSLKHP 720
Db 164 II-----PPPPPPPLPPPHITGP 183

RESULT 14
EMIL1_HUMAN
ID EMIL1_HUMAN STANDARD; PRT; 1016 AA.
AC Q9Y6C2; Q9G589; Q9UG76;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE EMILIN 1 precursor (Elastin microfibril interface-located protein 1) (Elastin microfibril interfac 1).
GN EMILIN1 OR EM1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=99287867; PubMed=10358019;
RA Doliana R., Mongiat M., Buccioti F., Giacomello E., Deutzmann R., Volpin D., Bressan G.M., Colombatti A.;
RT "EMILIN, a component of the elastic fiber and a new member of the C1q/tumor necrosis factor superfamily of proteins.";
RL J. Biol. Chem. 274:16773-16781(1999).
RN [2]
SEQUENCE FROM N.A.
RX MEDLINE=20092836; PubMed=10625608;
RA Doliana R., Canton A., Buccioti F., Mongiat M., Bonaudo P., Colombatti A.;
RT "Structure, chromosomal localization, and promoter analysis of the human elastin microfibril interface located protein (EMILIN) gene.";
RL J. Biol. Chem. 275:785-792(2000).
RN [3]
SEQUENCE OF 25-1016 FROM N.A.
RC TISSUE=Uterus;
RA Ottenwaelder B., Obermaier B., Mewes H.-W., Gassenhuber J., Wiemann S.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN [4]
SEQUENCE OF 434-1016 FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Vallalon D.K., Muzny K.C., Sodergren E.J., Lu X., Gibbs R.A., Fahy J., Helton A., Kettelman M., Madan A.C., Rodrigues S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.N., Krzyzanski M.I., Skalska U., Smalhus D.E., Schnerch A., Schein J.B., Jones S.J.M., Marra M.A.; "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";

PL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

RN [5]

RP TISSUE SPECIFICITY.

RX MEDLINE=21192249; PubMed=11278945;

RA Doliana R., Bot S., Mungiguerra G., Canton A., Cilli S.P., Colabattini A.;

RT "Isolation and characterization of EMILIN-2, a new component of the growing EMILINs family and a member of the EMI domain-containing superfamily.";

RT J. Biol. Chem. 276:12003-12011(2001).

RL -!- FUNCTION: May be responsible for anchoring smooth muscle cells to elastic fibers, and may be involved not only in the formation of the elastic fiber, but also in the processes that regulate vessel assembly. Has cell adhesive capacity.

CC -!- SUBUNIT: Homotrimer associated through a moderately stable interaction of the C-terminal globular Clq domains, allowing the nucleation of the triple helix and then a further quaternary assembly to higher order polymers via intermolecular disulfide bonds. Interacts with EMILIN2.

CC -!- SUBCELLULAR LOCATION: Extracellular matrix. Found mainly at the interface between amorphous elastin and microfibrils.

CC -!- TISSUE SPECIFICITY: Distributed in tissues where resilience and elastic recoil are prominent. Highest levels in the adult small intestine, aorta, lung, uterus, and appendix and in the fetal spleen, kidney, lung, and heart; intermediate expression was detected in adult liver, ovary, colon, stomach, lymph node and spleen; adult heart, bladder, prostate, adrenal gland, mammary gland, placenta and kidney showed low expression whereas a series of other adult tissues, including skeletal muscle and different regions of adult brain show no expression.

CC -!- MISCELLANEOUS: Its deposition precedes the appearance of elastin and is simultaneous with that of fibrillin 1.

CC -!- SIMILARITY: Contains 1 Clq domain.

CC -!- SIMILARITY: Contains 1 collagenous domain.

CC -!- SIMILARITY: Contains 1 EMI domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; AF088916; RAD42161.1; -

DR EMBL; AF162780; RAFP25006.1; -

DR EMBL; AL050138; CAB43287.1; -

DR EMBL; BC009947; AAH09947.1; -

DR PIR; T08772; T08772.

DR MIM; 130660; -

DR GO; GO:000515; F.protein binding; IPI.

DR GO; GO:0007155; P.cell adhesion; TAS.

DR InterPro; IPR001073; Clq.

DR InterPro; IPR008160; Collagen.

DR InterPro; IPR000577; FGGY kin.

DR InterPro; IPR008983; TNF_like.

DR Pfam; PF00386; Clq; 1.

DR Pfam; PF01391; Collagen; 1.

KW Cell adhesion; Extracellular matrix; Signal; Coiled coil; Collagen; Glycoprotein.

FT	SIGNAL	1	21	POTENTIAL.
FT CHAIN	22	1016	EMI.	EMI.
FT DOMAIN	57	129	COILED COIL (POTENTIAL).	COILED COIL (POTENTIAL).
FT DOMAIN	216	256	COILED COIL (POTENTIAL).	COILED COIL (POTENTIAL).
FT DOMAIN	356	420	COILED COIL (POTENTIAL).	COILED COIL (POTENTIAL).
FT DOMAIN	576	603	COILED COIL (POTENTIAL).	COILED COIL (POTENTIAL).
FT DOMAIN	685	752	COILED COIL (POTENTIAL).	COILED COIL (POTENTIAL).
FT DOMAIN	835	857	COLLAGEN.	COLLAGEN.
FT DOMAIN	814	864	C10.	C10.
FT DOMAIN	872	1007	N-LINKED (GLCNAC. . .) (POTENTIAL).	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD	154	154	N-LINKED (GLCNAC. . .) (POTENTIAL).	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD	415	415	N-LINKED (GLCNAC. . .) (POTENTIAL).	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD	455	455	N-LINKED (GLCNAC. . .) (POTENTIAL).	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD	561	561	N-LINKED (GLCNAC. . .) (POTENTIAL).	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD	658	658	N-LINKED (GLCNAC. . .) (POTENTIAL).	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD	766	766	N-LINKED (GLCNAC. . .) (POTENTIAL).	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD	794	794	PP -> TR (IN REF. 3).	PP -> TR (IN REF. 3).
FT CONFLICT	25	26	GLS -> HEG (IN REF. 4).	GLS -> HEG (IN REF. 4).
FT CONFLICT	434	436	6CF30238DD0EE26 CRC64;	6CF30238DD0EE26 CRC64;
SQ SEQUENCE	1016 AA;	106694 MW;	3.6%; Score 135.5; DB 1; Length 1016;	3.6%; Score 135.5; DB 1; Length 1016;
Query Match	Best Local Similarity	21.2%;	Pred. No. 0.34;	Mismatches 184; Conservative 71; Mismatches 297; Indels 315; Gaps 43;
QY	47 WRGCRVHRGD-	---	PVRVHVGNPQPHAVGCDLSTQI-----QCQ 85	---
DB	118 WRCCQGGDDCAESPAPALGPASSTPRPARPNLSSAGSPISGLGEGPGESEK	177	---	---
QY	86 LOSRAHQHQQIDKELQWRTG-	---	---AENLYRATS---NNR-----VRETVALE 126	---
DB	178 VQLEEQV-QSLTKELQGLRVLQGLSRLAEVDQRAVETAFNQRQPADAAARPGVHET	236	---	---
QY	127 LSYNSNLIQLL-	---	---KEELEBSG---GYDPRGHGSEAVTPMPLGLKXETKELDWST 177	---
DB	237 LNEHQHQLLDTRVTHDQELHNNHGGSSGSRAPAPAPAPGPPSEELLRLQLEQ	296	---	---
QY	178 PLKELISVHF-GEQG-	---	---ASYEAEIRELEALQAMRTPSR--- 213	---
DB	297 RLQESCVCLAGLDGFRQQQEDRRLAMEKLLASVEERQHLAAGVRRPPOCCSP	356	---	---
QY	214	---	---NESGLEL-----LTAYNQLCFLDARF-- 235	---
DB	357 ELGRRLAELERLDVAGSVTVLSGRRTTELGAAGCGHPGYTSLASRLSELDRENS	416	---	---
QY	236	---	---SLGLFFHWYDSLTVGPAQORALAFKGVSLFNIGALHTQIGARQDR 286	---
DB	417 TLGPSEEQEESWPAGPGGLSHW-	---	---LPAARGLE-QLGGLLANVSG---ELGRLD- 465	---
QY	287 SCTGARAMEAFQRAAGAFSLLENFHSAPDMSAAS--LCALEQLMMAQAQECVPEG	344	---	---
DB	466	---	---LLE-EQVAGAMQACQLCSGAPGEQDSQVSEILSALERRVL----- 505	---
QY	345 LSPASMAPQDCLAQLRL-	---	---AQEAQAQAAE--YRLVHRTMAQPPVHDVYVVS 392	---
DB	506	---	---DSEGQRLVSGSLHVTVEAAGEARQATLEGQEVGVGLQDRVDAQDQETAEE 555	---
QY	393 WTALVHVKAEPFRL-	---	---AHYHVAMALCDGSPATEGEL-----PTHEQVFLQPP 438	---
DB	556 FTLRNLTAARLQGLLEGLQAHDGCGACGQVQBELRLDRDGVRCSCF-	---	---LLEP 609	---
QY	439	---	---TSKPRGPVLPQELERERQLGKAHLKRAILQGEERLHALCRVREYDILLR 490	---
DB	610 RGPAGPGVGGPSRGFL-	---	---DGFSVFGSGSGSALQALQGESEVILSF 554	---
QY	491 AVISOTLQSLAKYAELEDREDDFCBAEAAPDIQPKTHQKPEARMPLRSCKGKPDIFHRIG	550	---	---
DB	655 SLSNDSLN-	---	---ELQTTVEGQGLADLAD-LG 679	---
QY	551 PLG--VFSANRWRLVGPVHLTRGEGGF-	---	---GLTLRGDSPVL-----IAA 590	---
DB	680 ATKDRIISEINRIQQEATBATESEERFRGLEEGQAQAGCCPSLEGLRGLEGVCEKLDLT	739	---	---

```

QY 591 VPGSQAAGLKEGDYIVSNGQPCRW---RHAENVTELKAA-----GEAGASLOV 640
Db 740 VAGGLQ---GLREG-----LSRHVAGIWAGLRENTTSQQAALLLEKLVGGQAGLGRRL 790
QY 641 VSLPSSRLPSLGR-----RPVLGPRGLLRSQREHGCKTPASTWASPRPLLNWSRKAQ 695
Db 791 GAL--NSSLQLLEDRLHQLSLKDLTGAG-----EAGPPGPGGLQGGPGAGPPGPGK 842
QY 696 QGKTG--GCPCP-----CAPVKPAP 713
Db 843 DQGGPIGPPGPGQEGQGVGAPAAPVP 869

RESULT 15
ARHC HUMAN
ID ARHC HUMAN STANDARD; PRT; 1544 AA.
AC Q9NZN5; O15086;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Rho guanine nucleotide exchange factor 12 (Leukemia-associated
DE RhoGEF).
GN ARHGEF12 OR LARG OR KIAA0382.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), AND CHROMOSOMAL TRANSLOCATION.
RC TISSUE=Prostate;
RX MEDLINE=20160919; PubMed=10681437;
RA Kourlas P.J., Strout M.P., Becknell B., Veronese M.L., Croce C.M.,
RA Theil K.S., Krahe R., Ruutu T., Knuttila S., Bloomfield C.D.,
RA Calligaris M.A.;
RT "Identification of a gene at 11q23 encoding a guanine nucleotide
RT exchange factor: Evidence for its fusion with MLL in acute myeloid
RT leukemia.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:2145-2150(2000).
RN [2]
RP SEQUENCE OF 795-1544 FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=97349984; PubMed=9205841;
RA Nagase T., Ishikawa K.-I., Nakajima D., Chira M., Seki N.,
RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.,
RT "Prediction of the coding sequences of unidentified human genes. VII.
RT The complete sequences of 100 new cDNA clones from brain which can
RT code for large proteins in vitro.";
RL DNA Res. 4:141-150(1997).
RN [3]
RP FUNCTION, AND INTERACTION WITH RHOA; GNA12 AND GNA13.
RX MEDLINE=20547532; PubMed=11094164;
RA Fukuhara S., Chikumi H., Gutkind J.S.;
RT "Leukemia-associated Rho guanine nucleotide exchange factor (LARG)
RT links heterotrimeric G proteins of the G(12) family to Rho.";
RL FEBS Lett. 485:183-188(2000).
RN [4]
RP INTERACTION WITH IGFIR.
RX MEDLINE=21581903; PubMed=11724822;
RA Taya S., Inagaki N., Sengiku H., Makino H., Iwamatsu A., Urakawa I.,
RA Nagao K., Kataoka S., Kaibuchi K.;
RT "Direct interaction of insulin-like growth factor-1 receptor with
RT leukemia-associated RhoGEF.";
RL J. Cell Biol. 155:809-820(2001).
RN [5]
RP INTERACTION WITH PLXNB1 AND PLXNB2.
RX MEDLINE=22259992; PubMed=12372594;
RA Driessens M.H., Olivo C., Nagata K., Inagaki M., Collard J.G.;
RT "B plexins activate Rho through PDZ-RhoGEF.";
RL FEBS Lett. 529:168-172(2002).
RN [6]
RP INTERACTION WITH PLXNB1 AND PLXNB2.
RX MEDLINE=22301907; PubMed=12183458;
RA Perrot V., Vazquez-Prado J., Gutkind J.S.;

```

```

RT "Plexin B regulates Rho through the guanine nucleotide exchange
RT factors leukemia-associated Rho GEF (LARG) and PDZ-RhoGEF.";
RL J. Biol. Chem. 277:43115-43120(2002).
CC -!- FUNCTION: May play a role in the regulation of RhoA GTPase by
CC guanine nucleotide-binding alpha-12 (GNA12) and alpha-13 (GNA13).
CC Acts as guanine nucleotide exchange factor (GEF) for RhoA GTPase
CC and may act as GTPase activating protein (GAP) for GNA12 and
CC GNA13.
CC -!- SUBUNIT: Interacts with GNA12 and GNA13, probably through the RGS-
CC like domain. Interacts with RHOA, PLXNB1 AND PLXNB2. Interacts
CC through its PDZ domain with IGFIR beta subunit.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic. Translocated to the membrane
CC upon stimulation (Probable).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q9NZN5-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q9NZN5-2; Sequence=VSP_008131;
CC -!- TISSUE SPECIFICITY: Ubiquitously expressed. Isoform 2 is found in
CC jejunum and testis.
CC -!- DISEASE: Involved in acute leukemia by a chromosomal translocation
CC t(11;11)(q23;q23) that involves MLL and ARHGEF12.
CC -!- SIMILARITY: Contains 1 DBL-homology (DH) domain.
CC -!- SIMILARITY: Contains 1 PDZ/DHR domain.
CC -!- SIMILARITY: Contains 1 PH domain.
CC -!- SIMILARITY: Contains 1 RGS(L) (RGS-like) domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; AF180681; AAF36817.1; -.
DR EMBL; AB002380; BAA20836.1; -.
DR HSSP; Q12923; 3PDZ.
DR Genew; HGNC:14193; ARHGEF12.
DR MIM; 604763; -.
DR InterPro; IPR001331; GDS_CDC24.
DR InterPro; IPR001478; PDZ.
DR InterPro; IPR001849; PH.
DR InterPro; IPR000342; Regl_Gpotein.
DR InterPro; IPR000219; RhoGEF.
DR Pfam; PF00595; PDZ; 1.
DR Pfam; PF00621; RhoGEF; 1.
DR SMART; SM00228; PDZ; 1.
DR SMART; SM00233; PH; 1.
DR SMART; SM00325; RhoGEF; 1.
DR PROSITE; PS00741; DH_1; FALSE_NEG.
DR PROSITE; PS50010; DH_2; 1.
DR PROSITE; PS50106; PDZ; 1.
DR PROSITE; PS50003; PH_DOMAIN; 1.
KW Guanine-nucleotide releasing factor; GTPase activation; Coiled coil;
KW Proto-oncogene; Chromosomal translocation; Alternative splicing.
FT DOMAIN 72 151
FT DOMAIN 367 558
FT DOMAIN 787 977
FT DOMAIN 1019 1132
FT DOMAIN 194 262
FT SITE 308 309
FT VARSPPLIC 48 66
FT Missing (in isoform 2).
FT CONFLICT 973 973 /FTid=VSP_008131.
FT Y -> F (IN REF. 2).
SQ SEQUENCE 1544 AA; 173231 MW; 0B7E319CF7C7A224 CRC64;

Query Match 3.5%; Score 132; DB 1; Length 1544;
Best Local Similarity 30.3%; Pred. No. 1;
Matches 36; Conservative 21; Mismatches 50; Indels 12; Gaps 3;

```

[illegible]

Search completed: July 13, 2004, 12:05:10
Job time : 28 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 13, 2004, 11:29:06 ; Search time 34 Seconds
(without alignments)

2045.485 Million cell updates/sec

Title: US-10-697-266-2

Perfect score: 3774

Sequence: 1 MLEERPDGAGAESPRLQ.....QPCAPVKPAPPSLKHPGWP 723

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 78:*

1: PIR1:*

2: PIR2:*

3: PIR3:*

4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	289	7.7	816	B86285	hypothetical prote
2	285.5	7.6	775	T37837	probable signal tr
3	284.5	7.5	882	S41034	hypothetical prote
4	232	6.1	661	S67177	hypothetical prote
5	202	5.4	844	S61104	BRO1 protein - yea
6	197.5	5.2	1400	T31555	hypothetical prote
7	155.5	4.1	942	JC2129	protein kinase PKN
8	141.5	3.7	358	I46532	protein co-factor
9	141	3.7	1734	A54602	microtubule-associ
10	139.5	3.7	943	T34847	probable transcrip
11	139.5	3.7	946	JC2130	protein kinase (BC
12	138.5	3.7	450	G01158	tyrosine kinase ac
13	135.5	3.6	992	T08772	hypothetical prote
14	135.5	3.6	2559	T09144	probable guanine n
15	130.5	3.5	1494	T14355	protein-tyrosine-p
16	130	3.4	1387	JC5502	G-protein signalin
17	129	3.4	1161	G81186	conserved hypothet
18	129	3.4	1289	B70748	probable smc prote
19	127.5	3.4	4684	A59404	plectin [imported]
20	127	3.4	1386	T00257	hypothetical prote
21	126.5	3.4	4574	G02520	plectin - human
22	126	3.3	4687	A39638	plectin - rat
23	125.5	3.3	2450	S71625	protein-tyrosine-p
24	124.5	3.3	1046	T42734	cytoplasmic linker
25	124	3.3	481	A56429	I-kappa-B-related
26	124	3.3	701	S62460	hypothetical prote
27	123.5	3.3	790	F75372	c-di-GMP phosphodi
28	123.5	3.3	10223	T30225	polyketide synthas
29	122.5	3.2	455	A55050	enigma - human

ALIGNMENTS

RESULT 1

B86285

hypothetical protein F9L1.7 [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana [mouse-ear cross]

C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Dec-2001

C;Accession: B86285

R;Theologos, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.; C.A.; Li, J.H.; Li, Y.; Liu, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziali, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I.

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A;Reference number: A86141; MUID:21016719; PMID:11130712

A;Accession: B86285

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-816 <STO>

A;Cross-References: GB:AE005172; NID:G5103812; PIDN:AAD39642.1; GSPDB:GN00141

C;Genetics:

A;Map position: 1

Query Match 7.7%; Score 289; DB 2; Length 816;

Best Local Similarity 23.3%; Pred. No. 4.9e-12; Mismatches 182; Indels 192; Gaps 24;

Matches 140; Conservative 88; Mismatches 182; Indels 192; Gaps 24;

QY 155 SEAVTPMIPPLGLKRTKELDWSTPLKELISVHFGEDGASYAEI--RELEALRQAMRTPS 212

Db 3 SSSLSNLMIAIHEKXTSSVDLYRLNRYVTFYSE---REAQLDDDLLETLKQ-LRSDI 57

QY 213 RNESG-----LELTATYNQLCFLDARF-LTPARS--LGLFFHWYDSLTGV-PAQORAL 262

Db 58 ERVSDPSPAARDLLISYKVLCLVETFPPIGFDKHVNAVSVYVWYAFKQKHKAQQNI 117

QY 263 APEKGSVLFNIGALHTQIGARQDRSCRGARRAMEAFORAGAFSLRENFS-----HA 316

Db 118 HLEKAALVFNLGASTYSQLGLGHDRITVDGRRQASHAFMAAGAFALHNRDNESIKATIGES 177

QY 317 PPSDMSAASLCALQELMMAQAQECVFE-----GLSPASMAPQDCIAQLRLAQEAQVAA 371

Db 178 TTVDVSVECVGMLERLWMAQAQECVFENTIAKGSIPGVS-----AKIAR 221

QY 372 EYRLVHRTWAQPPVHDYVPVSWTALVHVKAIEYFRSLAHVHMALCDGSPATEGELPTHE 431

Db 222 Q-----LKAALFYGEACFRYKEL-----HE 242

QY 432 QVFLQPPTSSKRGFPVLPQLEERFQLGKHLKRAILGQEEALRLHALCVLRVDLLRA 491

Db 243 K-----EETAE---ETALRSAGRLAEAKSKSRGAPAQLEAMNTLES 283

Qy	492	VISOTLQRS	LAKYAEL	DREDDP	CEAAEAP	DIQ----	PKTHQK	PEARP	RLSQG	KGPDI	546
		:	:	:	:	:	:	:	:	:	:
Db	284	SINGL	DRAV----	KENDRV	ILMRVP	SPSSLP	FAF	SMVK	PMNMT	DIID	ASKE
		:	:	:	:	:	:	:	:	:	:
Qy	547	HLRGL	SVFS	AKRW----	RLVGP	VHLTR	GEGG	FGT	LT	RG-	DSPV
		:	:	:	:	:	:	:	:	:	:
Db	337	SILVP	DSSAK	ALRYTE	VMVD	VIRTQ	AE	RLQ	QAS	ELTR--	VLK
		:	:	:	:	:	:	:	:	:	:
Qy	593	PGSQAA	AGLKE	GDIY	SVNS	GOPC----					616
		:	:	:	:	:	:	:	:	:	:
Db	390	DGNSAL	PVDL	KED	VEAVQ	SGG	FAG	LEA	ELQ	QLRDL	KRVN
		:	:	:	:	:	:	:	:	:	:
Qy	617	-----	RWR-	HA	EVTE----	-----	-----	-----	-----	-----	-----
		:	:	:	:	:	:	:	:	:	:
Db	450	FRSQ	FGTR	TRPQ	SSTL	TKNLQ	DLR	NR	FAAN	LKQ	AGES
		:	:	:	:	:	:	:	:	:	:
Qy	657	PV	658								
		:	:								
Db	509	PI	510								
		:	:								

RESULT 2
T37837
probable signal transduction pathway protein - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: T37837
R:Murphy, L.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
submitted to the EMBL Data Library, September 1997
A:Reference number: Z21749
A:Accession: T37837
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-775 <MUR>
A:Cross-references: EMBL:Z99162; PIDN:CA16216.1; GSPDB:GN00066; SPDB:SPAC17G6.05C
A:Experimental source: strain 972h; cosmid cl7G6
C:Genetics:
A:Gene: SPDB:SPAC17G6.05C
A:Map position: 1

Query Match	7.6%	Score 285.5	DB 2	Length 775
Best Local Similarity	25.3%	Pred. No. 7.9e-12		
Matches 112	Conservative 60	Mismatches 178	Indels 93	Gaps 15
QY	156	EAVTVMPILGKETKELDWSTPELKITLSVHFGDGSAYAEIRELEALRQAMRTPSRNE	215	
DB	2	EKLATPFYILNKKETHKSHDWEPTTFVSR.IYG.NSVDVEDQIKAFNTLRENAADVDVT	60	
QY	216	SGLELLTAYYNQLCFELDARFETPARSLGLFFHWDVSLTGVP----	AQRALAPKSGVLF	271
DB	61	AGKDIYGVYGQLDVLSFRFTGGNGINISFEWSILD--PDADFVKOSSIAFAKASYLF	118	
QY	272	NIGALHTQIGAFQDRSCT--EGARAMEAFORAAAGAFSILLRENFHAPSDMSAASLCALE	330	
DB	119	NLVSLTSMNAWHASAYTVDYKAAANCLQCASGIAKLLRESFIHAPGDLDSDNFLLGIY	178	
QY	331	QLMWAQAECEVFEGLSPSPASMAPQCLAQRLRAQAAQVAABYRLVHRTMAQPPVHDYVP	390	
DB	179	NLFLOQAQECVILGHMSFSAS---DSNMNYSLAAKIIASSAA-----	TLYDSCVHAPES	227
QY	391	V-----SWTALVHVKAAYFRSLAHYHVAMALCDGSPATEGELPTHEQVFLQPTTSSKPR	444	
DB	228	MEPACNPNFIRLASAKKAALLEGFSYFWARA-----		258
QY	445	GPVLPOELFERRQ---LGKAHLKRAILGOEEALRLHALCVLIREVDLLR-----	AVI	493
DB	259	-----QLEKSQGLAIGVLYQAOKSLTSSAOKL-----	FNIGIKLSTDFTHKPSLSDPQTI	308
QY	494	SOTLQRSIA---KYAELDR-----DDFCBAEAPDIQKTHOKPAPMRPLRSQK-----		541
DB	309	STFIKSSLSHLTKTAEKONDFVHDLVVKVEVSLPKISP-LQALPPLPLEKUYGSDGFT	367	

Qy	542	-----GDPFPHRLGPLSVPSAKN	559
Db	368	AKKIVGGDLFKAFVPSAVITASS	390
 RESULT 3			
S41034			
hypothetical protein R10E12.1 - Caenorhabditis elegans			
C;Species: Caenorhabditis elegans			
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Sep-1997			
A;Accession: S41034			
R;Smith, A.			
submitted to the EMBL Data Library, January 1994			
A;Reference number: S41034			
A;Accession: S41034			
A;Status: preliminary			
A;Molecule type: DNA			
A;Residues: 1-882 <SMI>			
A;Cross-references: EMBL:Z29561; NID:g450897; PID:g450898			
C;Genetics:			
A;Introns:	72/2;	324/3;	359/1; 400/2; 722/1; 743/1; 782/1; 818/1
 Query Match 7.5%; Score 284.5; DB 2; Length 882;			
Best Local Similarity 22.8%; Pred. No. 1.1e-11;			
Matches 132; Conservative 98; Mismatches 209; Indels 139; Gaps 22;			
Qy	167	LKETKELDWSTPLKELIS--VHFGEADGAEAE-IRELEALR-QAMTTP-SRNESGLELL	221
Db	11	LKSINEVDLVKPLTSYTDNVNTSDNNRSDVAAEVQLNKLRSKACCOPLDKHQSAIDLVL	70
Qy	222	TAYYNQLCFLDARELTLPARSLGLFFHWYDSL-TGVPAQQRAL-----APEKGSVLENIG	274
Db	71	TRYDYLVAIEINKIIISATQNPPVVKWDAPDKGSLFSRASLSLDSGSPERAVALFNIG	130
Qy	275	ALHTQIGARODRSCTEGARAMEAFORAAAGFSLRLR---NFSSHAPSFDMSAASLCAL	330
Db	131	SLSMQIGAAQOFHTDDIKYSAKLFQSAGVFARLDVULGMWQOEETPDLMPTDLAALS	190
Qy	331	QLMMAQAQCEVFEGLSPPASMAPQDCIAQLRLAQEAACVAAYKLVRHTWAQPVDHVVP	390
Db	191	ALMTAQAEAIY-----IKGHDKMKATSMVKISAQVAEFYSEAOKMWSKDIVRGILWD	243
Qy	391	VSWTALVHVKAIFYFRSLAHVHAVNALCDGPSPATGEGELPTHEQVFLQPTSSKPRGPVL	450
Db	244	KDWSAIVSGKNLAYCALAQYHQSS-EVCG-----	270
Qy	451	ELEERRQLGKAHLRAILGOBEALRHLCRVLREVLLRAVISQ-----TLQRS Lak-	503
Db	271	--EARIGE-----QLSRLESKLFTAQKYLPDITGIWDIYPSVSKA	313
Qy	504	YAELDREDDFCEAAEAPDIQPKTHOKPEARMPRLISOQKGK-----DIFHRLGPLSVPSA	557
Db	314	HAAAKKNDFTIYEKVSDFR-TLPTLPKAVLAKEPTPMQTETPTSPFRDMFAVLVPVQHNA	372
Qy	558	KNRW-----RLVCPVHLTRCEGGFGILTLCGDSPVLTAAVIPEGQAAGAALKE	604
Db	373	MQSYDARKAELVNMTVMREATQMNG-----VLASLNLPAL-----	411
Qy	605	GDIYVSUNGQPCRWWRHAENVTELKAA--GEAGASLOWVSLPSRSSLPGLDRRRFVLIGP	662
Db	412	-DDVTSTETLP-----ESLKLSAKLXKNGSSSEIMRLF--SELPTLYQRNEIDIITE	460
Qy	663	RGLRSQREHOCKTPTASTWASPRLNNSRKAAQOQKTG	700
Db	461	TSRIINLESKEDDTMRKOLGF-----KWTRMSSPQLTG	493

RESULT 4
S67177
hypothetical protein YOR275c - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein O5450
C:Species: Saccharomyces cerevisiae
C:date: 12-Jul-1996 #sequence revision 12-Jul-1996 #text change 19-Apr-2002

```

315 HAPSPDMSAASLCALEQLMMAQAQECVFE----- 343 QY
      |||:::|||||||: Db
153 FYPSVDLDANVISFYKYVMIAQEELCQLVKSLDDNRSATVIAKLSWLQFAYDSAAKIVD 212
      |||:::|||||||: QY
344 --GLSPPPSAMAPQCCLAURLAQEAQAQVAEYRLVHRTMQAPPVHPDPVFSWTALVHVKA 401

```

```
Db 213 EWSNIPESVQRRTAIKLDLKQNAEKENDF-----IYHDRVPKQEDAIQVOK 262
Qy 402 E-----YFRSLAHVHVMALCDGSPATEGELPHEQVFLQPTSSKRGPGVLPQLEER 455
Db 263 DGGAMCKVKTLSPDLDPDPSVC-GCDLFGKLLPTFVQDAVKYSDDK-----DQALRE- 314
Qy 456 ROLGKAHLKRAILGOEEALRLHALCRVLRREVLLRAVISOTLQSLAKYAELEDREDFCE 515
Db 315 -----IKECVRSYDEHLNVLQ-----QAEDFKURFMLNEG-KRSREAWFELS-EDLMKR 362
Qy 516 AAEA---PDIQPKTHQK---PEARPRLSGQKGPDIHRLGPLSVFSAKNRWLVGPVHLT 570
Db 363 NADMTSYEDCPNLDKMGESSDTRVAEAKINTLLSKLRAIDLQKSS-----I 411
Qy 571 RGEFGGLTLRGDSPVLIAVIPSQAAAAGLKEGD-YIVSVNGQPCRWWRHAEVVTBLK 629
Db 412 RSDEGF-----ILIQKEL-----ERLAHLEQAKAHNVSLN-----K 443
Qy 630 AAGRAGASIQVVSL 643
Db 444 ALAQHSANLQLTL 457

RESULT 7
JC2129
protein kinase PKN (EC 2.7.1.-) - human
N:Alternate names: protein-kinase-C-related kinase PRK1
C:Species: Homo sapiens (man)
C:Date: 30-Sep-1993 #sequence revision 20-Aug-1994 #text_change 17-Nov-2000
C:Accession: JC2129; S67526; S51020; I53327
R:Mukai, H.; Ono, Y.
Biochem. Biophys. Res. Commun. 199, 897-904, 1994
A:Title: A novel protein kinase with leucine zipper-like sequences: Its catalytic domain
A:Reference number: JC2129; MUID:94183274; PMID:8135837
A:Accession: JC2129
A:Molecule type: mRNA
A:Residues: 1-942 <MUK>
A:Cross-references: GB:D26181; NID:9473712; PIDN:BA005169.1; PID:9825505
R:Palmer, R.H.; Riddien, J.; Parker, P.J.
Eur. J. Biochem. 227, 344-351, 1995
A:Title: Cloning and expression patterns of two members of a novel protein-kinase-C-related
A:Reference number: I53327; MUID:95154310; PMID:7851406
A:Accession: S67526
A:Molecule type: mRNA
A:Residues: 1-190,'D',192-942 <PAL>
A:Cross-references: EMBL:S75546; NID:9914097; PIDN:AAB33345.1; PID:9914098
A:Experimental source: fetal brain
R:Palmer, R.H.; Riddien, J.; Parker, P.J.
FEBS Lett. 356, 5-8, 1994
A:Title: Identification of multiple, novel, protein kinase C-related gene products.
A:Reference number: S51020; MUID:95080426; PMID:7988719
A:Accession: S51020
A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra
A:Molecule type: mRNA
A:Residues: 700-799,'A' <PA2>
C:Comment: This protein has a protein kinase domain related to protein kinase C.
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo
C:Keywords: ATP; leucine zipper; phosphotransferase; protein kinase
F:39-66/Region: basic
F:70-287/Region: leucine zipper motif
F:613-874/Domain: protein kinase homology <KIN>
F:621-629/Region: protein kinase ATP-binding motif
F:644/Active site: Lys #status predicted

Query Match 4.1%; Score 155.5; DB 2; Length 942;
Best Local Similarity 21.4%; Pred. No. 0.012;
Matches 159; Conservative 76; Mismatches 260; Indels 247; Gaps 36;

Qy 69 QLHAVGCDLSITQCQQLQSRRAQIHQIQDKELQMTGAENLYRATSN-NRVRETVALEL 127
Db 18 QLGLAGADLAAPGVQQLERERURRETKELKKEGAEENLRATDTLGRSLGPEVLL 77
```

```
Qy 128 SYVNSNQLLKELEELSGGV---DP-GRHSEAVTVPMIPLGLKETKELDMSTPLKELI 183
Db 78 RGSRRDLLHQQLQELHAHVLPDPAATHDG-----PQSP----- 113
Qy 184 SVHFGEDGASVEA-EIELEALROAMRTSPSNESGLELLTAYYNQLCFDARFLTPARSL 242
Db 114 ---GAGPTCSATNLSRVAGLEKQALIELVKQGAENMIQTYNSGTCKRKLTLTAQOM 169
Qy 243 GLFFHWYDSLT-----GVPAQORALAFEKGSVLPFNIGALHTQIGA----- 282
Db 170 -----LQDSKTKIDILRMQLRRALQAGLENAQAPDDTQSGP--DLGAVELRIELRHHF 222
Qy 283 QDRSCCTEGARRAMEAFQRAAGAFSLLENFSPHAPSMDMSAASLCALQLMMAQAQECVF 342
Db 223 RVEHAVAEGAKNVLRL-----LSAAKAPDRKAVSEAAQEKLTESNQKLGLLR 268
Qy 343 EGLSPPSMADQDCLAQLRLAQEAQAQVAAEYELVHRTMAQPPVHDY-----VPVSWTALVH 398
Db 269 EALERRIGELPAD-HPKGRLLREELAAASSAFTRLAGPPPATHYSTLCKPAPITGTILE 327
Qy 399 VKAEYFRSLAH---YHVAMALCDGSPATEGELPHEQVFLQPP----- 438
Db 328 VRVVCRRDLPETIPWNTPSM--GGPGT---PDSRPFFLSRPARGLYSRSGSLSGRSSL 381
Qy 439 -----TSSKPRGP-----VLPQLEERRQLGKAHLKRAILQO 470
Db 382 KAEANTSEVSTVLKLDNTVVGQTSWKPCGPNAMDQSTLELERARE-----LELAVFWR 436
Qy 471 BEALRLHALC--RVLRVDLLRAVISOTLQSLAKYAELEDREDDFCEAAAPDIQPKTHQ 528
Db 437 DQ---RGLCALFKLEDFL-----DNERHEVQLDMEPQOCLVAEV-----TFR 477
Qy 529 KPE-ARMPRL-----SQCKGPDIFHRLGPLSVFSAKNRWLVGPVHLTR----- 571
Db 478 NPVIERIFLRQKKIFSKQOQKA---FQARQNMIDVA--TW-----VRLRLRLIPNAT 527
Qy 572 GEGFGLTLRGDSPVLIAAVIPGSOAAAGLKEGDIYIVSVNGQPCRWWRHAEVVTELKAA 631
Db 528 GTGTF-----SP---GASPGSEARTTG-----DISVE-----KMLIG 556
Qy 632 GEAGASLQVSVLLPSSRLPSLGRDPRVLLGPRGLLRSQREHCKTCKPASTWASPRLLNWS 691
Db 557 TDSDSPOKSRDRDPSSPSSLS--PI-----QESTAPELPSETQETFGPAL--- 601
Qy 692 RKAQOQKGTGCGPCQPCAPVKPAP 713
Db 602 -----CSPLRKSP 609

RESULT 8
I46532
protein co-factor - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 14-Feb-1997 #sequence_revision 14-Feb-1997 #text_change 05-Nov-1999
C:Accession: I46532
R:Weinman, E.J.; Stepien, D.; Wang, Y.; Shenolikar, S.
J. Clin. Invest. 95, 2143-2149, 1995
A:Title: Characterization of a protein cofactor that mediates protein kinase A regulation
A:Reference number: I46532; MUID:95256437; PMID:7738182
A:Accession: I46532
A:Status: preliminary; translated from GB/EMBL/DOBJ
A:Molecule type: mRNA
A:Residues: 1-358 <WEI>
A:Cross-references: EMBL:U19815; NID:9687674; PIDN:AAA80218.1; PID:g687675
C:Superfamily: GLGF domain homology
F:19-90/Domain: GLGF domain homology <GLG2>

Query Match 3.7%; Score 141.5; DB 2; Length 358;
Best Local Similarity 24.0%; Pred. No. 0.028;
Matches 89; Conservative 31; Mismatches 96; Indels 155; Gaps 16;

Qy 419 GSPATG-----ELPTHEQVFLQPTSSKPRGP-----VLPQLEER 454
```

Db 45 GSPAERAGLAGLRLVBNVNGENVEKETHQVW-----SRIRALNAVRLVVDPTDEQ 98
QY 455 RROLG---KAHLKRAILGQEAHLRLHLCVLRVLDLLRAVISOQLORSKAKYAELDRED 511
Db 99 FRKLGVOIRGELLRAQAGPEQAGPPAAP----- 126
QY 512 DFCEAAEAAPDIQKTHQK--PEAR--MPRLSQGKGPDI FHLRGLPLSVFSKAKNRWLVGPV 567
Db 127 --GEQPAGEGENEPEVEKSHPERELRPLCAMK----- 158
QY 568 HLTRGEGGFLTURGDS--VLIAAVIPGSOAAAAGLKEGDYIVSYNGQPCRWWRHAEV 624
Db 159 ---KGPNGYGNLHSDKSRGQFIRAVDPDS PAEASGLRQDRIVEVNGVCVEGKQGDV 215
QY 625 VTELKAGEAGASIQVW-----SLLPSR-----LP-----SLGDRRPVLLGP 662
Db 216 VTAIKAGGDE-AKLLVVVDKETDEFFKCKVVPSEHLNGPLPFPFTNGEIQKNNPETLAP 274
QY 663 R-----GLLR-----SQRHGGCKTPASTWASPRPLLNWS-----RK 693
Db 275 AAASESRPALARSASSDTSEELASQDSPKEDSTAPSTSSSSDPIIDFISLAVAKERA 334
QY 694 AQQCKTGGCPQ 704
Db 335 HQKRSSRRAPQ 345

RESULT 9

A54602

microtubule-associated serine/threonine protein kinase MAST205 - mouse

C:Species: Mus musculus (house mouse)

C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 24-Sep-1999

C:Accession: A54602

R:Walden, P.D.; Cowan, N.J.

Mol Cell. Biol. 13, 7625-7635, 1993

A:Title: A novel 205-kilodalton testis-specific serine/threonine protein kinase associated

A:Reference number: A54602; MUID:94067123; PMID:8246979

A:Accession: A54602

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-1734 <NAL>

A:Cross-references: GB:U02313; NID:9406057; PIDN:AAC04312.1; PID:9406058

C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo

C:Keywords: ATP

F:451-726/Domain: protein kinase homology <KIN>

F:459-467/Region: protein kinase ATP-binding motif

Query Match 3.7%; Score 141; DB 2; Length 1734;
Best Local Similarity 21.3%; Pred. No. 0.28;
Matches 135; Conservative 62; Mismatches 238; Indels 200; Gaps 26;

QY 161 PMIPLGKTK-----LDWSTPLX---ELISVHFGEGCASY----- 194
Db 708 PLERLTSSAYEVKQHPFPMGLDWTGLLRQKAEFIPQLESEDDTSYFDRSRBYHVDSE 767
QY 195 -EAEIRE---LEALROAMTP--SRNESGLELLTAYYNOLCFIDARFLPARSLGIFPHW 248
Db 768 DEEEVSDGCLIRQSSCSPPFSKYSSNE-----RLSLEER-RTP----- 809
QY 249 YDSLTVPAQQALAFKSGVLFNIGALHTQIGARQDRSCTEGARRAMEAFQAAAGFSL 308
Db 810 -----PPTKESLSEKEDHSDGLAGL-----KGRDRSWIGVSPILR--KRLSVSESS 855
QY 309 LRENFSHAPSDMSAASLCALQLMWAQAQCVFGLSPPMASPAQDCLAQLRLAQEAQA 368
Db 856 HTESDS---SPMTVRHRS-----GLPDGPHCFEETSSTP----- 888
QY 369 VAAEYRLVHRTMAQPPVHDYVPVSWTALVHVVAEYFRSLAHYHVAMALCDGSPATEGELP 428
Db 889 -----RKQQQEGIVLIIPSG-----EGSSRPVPRP 915
QY 429 THEQVFL--QPPTSSKPRGPVLQOELEERRQLGKAHLKRAILQOEALRHALC----- 480

Db 916 LERQLKLDEBPPOGSSRCCPALETROGTQQLABEATAKAISDLAVRRARHRLSGDSIE 975
QY 481 -RVLREVDLLRAVISOQLORSKAKYAELDREDDFCEAAEAAPDIQKTHQPEAR----- 533
Db 976 KRITRPUNKVIKASATALSL-----IPSEHACSPASPMSPHSOSSPSSRSDSFSR 1030
QY 534 --MPRLSQGKGPDI FHLRGLPLSVFSKAKNRWLVGPVHLTRGEGGFLTLR-----GDSP 585
Db 1031 DFLPALGSLRPPIIIHRAGK-----KYGFTLRAIRVYMGDTD 1067
QY 586 V-----LIIAAVIPGSOAAAAGLKEGDYIVSYNGQPCRWWRHAEVVTTELKAGAGASLQV 640
Db 1068 VYVTHMVMHVEGGPASEAGLFOGDLITHVNGEPVHGLVHTEVVELVLSGNGK-VSIST 1126
QY 641 VSLPSSRLPSLGDRRPVLGPRGLLRQSREHG-----CKTPASTWASPR 685
Db 1127 TPLENTS--IKVGPARKGSYKAWARRSKSKGQGESKRSSLFRKITQASLLHTR 1184
QY 686 PLLNWRKKAQKQKTG-GCPQPCAPVKP-APPSSLK 718
Db 1185 SLSSLNRLSSGSGSPGSPTHSHLSLSPRPPOGYR 1219

RESULT 10

T34847

probable transcription regulator - Streptomyces coelicolor

C:Species: Streptomyces coelicolor

C>Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 05-Nov-1999

C:Accession: T34847

R:Oliver, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, February 1999

A:Reference number: Z21559

A:Accession: T34847

A>Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-943 <OLI>

A:Cross-references: EMBL:AL035478; PIDN:CAE36601.1; GSPDB:GN00070; SCOEDB:SC2G5.14C

A:Experimental source: strain A3(2)

C:Genetics:

A:Gene: SCOEDB:SC2G5.14C

Query Match 3.7%; Score 139.5; DB 2; Length 943;
Best Local Similarity 22.3%; Pred. No. 0.15;
Matches 177; Conservative 62; Mismatches 253; Indels 301; Gaps 39;

QY 55 GEDPVRV-----HVGPMNQLHAGVCDLSLTIQCGQLQSR--AQIH 94
Db 166 GETPVRLCAVTEGTETDDAEYDHLRASPPDTRAVRLGPLTRAQVSAALLDGRGYSLH 225
QY 95 QQIDKELQMRGAEENLYRATSNRRVRETVAL-----ELSYVNSNLQLLKEELEELSG--- 146
Db 226 RSTVRDIHRTSGGNPLFALELGRALAEPAAPRPGEPVPVTSRLALVLSRLLEMLSGEAR 285
QY 147 -----GVDP-----GRHGEANTVPMPLGL-----KETELDWSTPLKELISVH 186
Db 286 RLLVASAGARTPALLRAAGRENAEETVQAALGLLAPDPEGPTLRFAPH--LIS--- 340
QY 187 FGEDGASY-EAEIRELEALROAMRTPSNESGLELLTAYYNOLCFIDARFLTPARSGLIF 245
Db 341 ---AALYAEAPAPERAAHAAALSTAADP-----TERARHLAL- 375
QY 246 FHWYDSLTVGP-----AQRALAFEGK--SVLFNIGAL---HT-----QIGA 282
Db 376 -----ATTGTDPEVAARLAEEAALARDGAPSAASLGLLAARHTPADGTPGPDERRUTA 430
QY 283 RQD-----RSCTEGAR-RAMEAFORAAGA----- 305
Db 431 ABDATTAGEQDLARDVAREVLTRATAPGERVRAWMVAITEAAGQATGEVDVFPQALADAG 490
QY 306 -----FSLLENFSHAPSDMSAASLCA-----LEOLMMA- 335
Db 491 DDPRLALVHYQLAWSRLVVGQDFAQGREAAH-----AARLAARAGERSTELLALAF 543


```
Db      297 SGLHLSPTAAARRRLEPCBSTSARHRWTGTGSKSSATSEPLPACLTGLTGLPLPHGPMAS 356
QY      684 PRLLNWSRAQOQKTCGCPQCPAPVKA-----PPSLKHFG 721
Db      357 ACPEL-----PQFWTGGWSCHCFEISPSFGEPPSCPCPPG 392

RESULT 13
T08772
hypothetical protein DKFZp586M121.1 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 13-Aug-1999
C;Accession: T08772
R;Ottenwaelder, B.; Obermaier, B.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, May 1999
A;Reference number: Z16474
A;Accession: T08772
A;Molecule type: mRNA
A;Residues: 1-992 <OTT>
A;Cross-references: EMBL:AL050138
A;Experimental source: adult uterus; clone DKFZp586M121
C;Genetics:
A;Note: DKFZp586M121.1

Query Match      3.6%; Score 135.5; DB 2; Length 992;
Best Local Similarity 21.2%; Pred. No. 0.31;
Matches 184; Conservative 71; Mismatches 297; Indels 315; Gaps 43;

QY      47 WRCRVHRGED-----PVRVHVGPMNPLHAGVCDLSTQI-----CCQG 85
Db      94 WRCCQGYGGDDCAESPAPALGPASSPTPLRLARPENLSSAGSLGGLGGPGGESEK 153
QY      86 LQSRRAQIHQIDKEIQMRTG-----AENLYRATS-----VRETVALE 126
Db      154 VQLEEQV-QSUTKELQGLRGVLQGLSLAEVDQRAVETAFNGRQCPADAAARPVGWHT 212
QY      127 LSYVSNLQLL-----KEELEELSG--GVDPGRHGSVAITVPMIPLGKTKELDWST 177
Db      213 LNEIQHQLQLLDFRVTHQOELGHNNHGGSSSGSRAPAPASPPGSEELLQLQLEQ 272
QY      178 PLKELISVHF-GDQG-----ASYAEIREEALROAMRTPSR-----213
Db      273 RLQESCSVLGLDGDFRQOQEDRERLRAMEKLLASVEERQRLHAGLVARRPPQCCSP 332
QY      214 -----NESGLEL-----LTAYYNQLCFLDARF--235
Db      333 ELGRRLAELERLDVVAGSVTVLSGRGTGLGGAQGGHPGYTSLASRLSLREDFNS 392
QY      236 -LTPAR-----SLGLFFHWYDSLITGVPAAQRLAFEFKSVLEFNGALHTQIGARQDR 286
Db      393 TLGPSEEQEESWFGAFGLSHW-----LPAARGRLLE-QLGGLLANVSG-----ELGGRLLD- 441
QY      287 SCTEGARRAMEAPQRAAGAFSLLENFNSHAPSPDMSAAS--LCALQLMMAQAQCEVFRG 344
Db      442 -----LLE--EQVAGAMQACGLQCSGAPGEQDSQVSEILLSALERRVL-----481
QY      345 LSPSPASMAPQDCIAQLRL-----AQEAQAQVAE--YRLVHRTMAQPPVHDVVPVS 392
Db      482 -----DSEGQRLVGSGLHTVEAAGEARQATLEGLQEVVGRVLRDVAQDETAAE 531
QY      393 WTALVHVKAERYFRSL-----AHVHVAMALCDGSPATEGEL-----PTEQVFLQPP 438
Db      532 FTLLRLNLTAARLQGLGLLAHQHDECGACGGVQBELGRLRGVERCSCP-----LLPP 585
QY      439 -----TSSPRGVPVLQPELEERRQLKAHLKRAILQGEAEALRLHALCRVIREVDLLR 490
Db      586 RGFAGPGVGGPSRGPL-----DGFSVFGSGSALQALQGLSEVILSF 630
QY      491 AVISQTLQRLAKYAELEDREDDFCENAEAPDTQPKHQPEARMPLRSQKGPDPFIRLIG 550
Db      631 SSLSNLSL-----ELQTTVEGQADLAD-LG 655
QY      551 PLS--VFSAKNRWRLVGPVHLTRGEGGF-----GLTLRGDSPVL-----INA 590
```

```
Db      656 ATKDRIISEINRLQOQATEHATESERFRGLREGQAQACQPSLEGRGLRGLEGVCERLDT 715
QY      591 VIPGQAAAAAGLKEGDYIVSVNGQPCRW-----RHAENVVTELKAA-----GEAGASLQV 640
Db      716 VAGGLQ-----GLREG-----LSRHVAGLWAGLRETNTTSQMAALLLEKLVGGQAGLGRRL 766
QY      641 VSLPSSRLPSLQDR-----RPVLLGFERGLLRSQRHCKTKTPASTWASPRPLLNWSRKAQ 695
Db      767 GAL--NSSLQLLEDRLHLQSLKDLTGPA-----EAGPPGPPGLQGPAGPGPGSPGK 818
QY      696 QGKTG--GCPQP-----CAPVKPAP 713
Db      819 DQEGPIGPPPGQEQGVGEGAPAAPVP 845

RESULT 14
T09144
probable guanine nucleotide exchange factor RhoGEF2 - fruit fly (Drosophila melanogaster)
N;Alternate names: Shar pei/DRHOGEF2
C;Species: Drosophila melanogaster
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 08-Sep-2002
C;Accession: T09144; T09223
R;Haecker, U.; Perrimon, N.
submitted to the EMBL Data Library, October 1997
A;Reference number: Z16586
A;Accession: T09144
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-2559 <HAE>
A;Cross-references: EMBL:AF031930; NID:g2687355; PID:g2687356
R;Barrett, K.; Leptin, M.; Settleman, J.
Cell 91, 905-915, 1997
A;Title: The Rho GTPase and a putative RhoGEF mediate a signaling pathway for the cell st
A;Reference number: Z16618; MUID:9808790; PMID:9428514
A;Accession: T09223
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-676,'L',678-837,'L',839-889,'S',891-1243,'D',1245-1358,'E',1360-1368,'R',13
A;Cross-references: EMBL:AF032870; NID:g2760367; PIDN:AAC38820.1; PID:g2760368
C;Genetics:
A;Gene: rhoGEF2
A;Cross-references: FlyBase:FBgn0023172
A;Map position: 2; 53F1-2
A;Note: orchestrates cell shape changes during gastrulation
C;Function:
A;Description: mediates actin rearrangements required for cell shape changes during gastr
C;Superfamily: rat Munc13-3 protein; protein kinase C zinc-binding repeat homology
C;Keywords: embryo; GTP exchange; signal transduction
F;1151-1200/Domain: protein kinase C zinc-binding repeat homology <KZN>

Query Match      3.6%; Score 135.5; DB 2; Length 2559;
Best Local Similarity 25.1%; Pred. No. 1.2;
Matches 50; Conservative 27; Mismatches 71; Indels 51; Gaps 6;

QY      483 LREVDLLRAVISQTLQRLSLAKYAELEDR--DDFCFAAEAPDIQPKHQKPEARMPLRSQ 540
Db      202 LNQSDNSNPVLQAPGERSLNLTLPLSRDLSSGHTQBSTTTPATPTST---PSLALPKNFQ- 257
QY      541 KGPDPFHLRLPLSVFSAKNRWFLVGPVHLT-----RGEQGFGLTLRGDSPVLIAAVIPG 596
Db      258 -----YLTUTVRKDSNGYGMKVGSDNPFVFVESVKPGA 290
QY      597 AAAAGLKEGDYIVSVNGQPCRWHRHAENVVTELKAAEA-----GASLQVSVLL 644
Db      291 RELIAGVAGDMLLRVNGHEVRLEKPTVVGLIKASTTVELAVKRSQKLTSPSSVSVT-- 348
QY      645 PSSRLPSLGDRRPVLGLGR 663
Db      349 PSTPILSGRDRTASITGPQ 367

RESULT 15
```

T14355
protein-tyrosine-phosphatase (EC 3.1.3.48) TD14 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C;Accession: T14355
R;Cao, L.; Zhang, L.; Ruiz-Lozano, P.; Yang, Q.; Chien, K.R.; Graham, R.M.; Zhou, M.
J. Biol. Chem. 273, 21077-21083, 1998
A;Title: A novel putative protein-tyrosine phosphatase contains a BRO1-like domain and s
A;Reference number: 218004; MUID:98361981; PMID:9694860
A;Accession: T14355
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1494 <CAO>
A;Cross-references: EMBL:AF077000; NID:g3598973; PID:g3598974; PIDN:AAC62959.1
A;Experimental source: brain
C;Genetics:
A;Gene: PTP-TD14
C;Function:
A;Description: may be involved in regulating Ha-ras-dependent cell growth
C;Keywords: phosphoric monoester hydrolase

Query Match 3.5%; Score 130.5; DB 2; Length 1494;
Best Local Similarity 23.5%; Pred. No. 1.2;
Matches 56; Conservative 36; Mismatches 87; Indels 59; Gaps 9;

QY 333 MMAQAQECVFEGLGPPASMAPODCLAQRLAEEAAQVAAEYRLVHRTMAQPPVHDY---V 389
Db 1 MLGQAQECLE-----KSMLE--DNEKSLVARISAQVVDYKACACRALENPTASLLGRI 53

QY 390 PVSMTALVHVKAERYFRSLAHYHVMALCDGSPATEGELPTHEQVFLQPPTSSKRGVPVLP 449
Db 54 QKDWKKLVQMKIYYFAVAHLHMG----- 77

QY 450 QELEERRQLGK--AHLKEALIGQEEALRLHALCRVIREVDLLRAVISQTLQESLAKYAEEL 507
Db 78 KOAEQQKFEGERVAYFQSGALDKLNEAIKL-----AKGQDPTVQDRLRFAMDVIGGKYN 132

QY 508 DREDDFC--EAAEAPD-IQPKTHQ---KPEARMPLISQKGPDI FHRIGPLSVFSAKN 559
Db 133 KKNDFIYHEAVPALDTLPQVKGAPLVKPLVNPPTDPAVTGPDIFAKLVPMAAHEASS 190

Search completed: July 13, 2004, 12:07:49
Job time : 37 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 13, 2004, 08:21:11 ; Search time 87 Seconds
(without alignments)

2348.067 Million cell updates/sec

Title: US-10-697-266-2

Perfect score: 3774

Sequence: 1 MILERRPDGAGAGEESPRLO.....QPCAPVKPAPPSLKHPGWP 723

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001s:*
5: Geneseq2002s:*
6: Geneseq2003as:*
7: Geneseq2003bs:*
8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3774	100.0	723	5	AAO18602 Human pro
2	3449.5	91.4	670	6	AB999400 Amino aci
3	2331	61.8	634	5	AAO18603 Human pro
4	2331	61.8	643	6	ABM32685 Human Rho
5	1401	37.1	272	6	ABR41761 Human DIT
6	1278	33.9	257	3	ABR43571 Human can
7	1278	33.9	257	3	AAU00869 Human can
8	1255	33.3	611	5	ABU65216 Human NOV
9	1249	33.3	686	4	AAU10192 Human pro
10	1249	33.1	686	5	ABG94709 Human GTP
11	1236	32.8	666	6	ABG68561 Human GTP
12	1195	31.7	686	5	ABU65215 Human NOV
13	1107	26.7	1111	4	ABG08369 Human NOV
14	972	25.8	556	5	ABG61936 Novel hum
15	817	21.6	648	4	ABG63879 Prostata
16	766	20.3	488	4	ABG64579 Drosophil
17	590.5	15.6	321	4	ABG06219 Human tra
18	413.5	11.0	110	3	ABG41991 Human ORF
19	394.5	10.5	311	4	ABG08364 Novel hum
20	380.5	10.1	939	6	ADA54327 Human kin
21	380.5	10.1	1636	5	AAE25081 Human kin
22	377.5	10.0	1636	5	AAE25081 Human kin
23	362.5	9.6	308	4	ABG06220 Novel hum
24	341	9.0	170	3	AAE41575 Human ORF
25	304.5	8.1	1583	3	AAE29662 Human tyr

ALIGNMENTS

RESULT 1

ID	AAO18602	standard; protein; 723 AA.
XX	AAO18602;	
XX		
DT	24-OCT-2002 (first entry)	
XX		
DE	Human protein kinase N family kinase.	
XX		
KW	Human; protein kinase N; kinase; PKN; eye retinoblastoma; bocio tumour;	
KW	placenta choriocarcinoma; germ cell; leukocyte; uterus tumour; cancer;	
KW	pre-B cell acute lymphoblastic leukaemia; Wilm's tumour; enzyme;	
KW	brain anaplastic oligodendroma; uterus endometrial adenocarcinoma;	
KW	cytostatic; gene therapy; chromosome 8.	
XX		
OS	Homo sapiens.	
XX		
PN	WO200261062-A2.	
XX		
PD	08-AUG-2002.	
XX		
PF	29-JAN-2002; 2002WO-US002152.	
XX		
PR	01-FEB-2001; 2001US-00773371.	
PR	07-MAY-2001; 2001US-00849334.	
XX		
PA	(PEKE) PE CORP NY.	
XX		
PI	Rusch D, Ketchum KA, Di Francesco V, Beasley EM;	
XX		
DR	WPI; 2002-608516/65.	
DR	N-PSDB; AAL48837, AAL48838.	
XX		
PT	New human kinase peptide and nucleic acid molecule, useful for treating	
PT	disorders associated with abnormal expression of kinase protein, e.g.	
PT	retinoblastoma, Wilm's tumor, in drug screening assays and	
PT	pharmacogenomic analysis.	
XX		
PS	Claim 1; Fig 2; 76pp; English.	
XX		
CC	The present invention provides the protein and coding sequences of a	
CC	novel human protein kinase N family protein. The sequences can be used in	
CC	the treatment of disorders associated with the absence of, inappropriate,	
CC	or unwanted expression of the protein, e.g. eye retinoblastoma, Wilm's	
CC	tumour, placenta choriocarcinomas, bocio tumours, pre-B cell acute	
CC	lymphoblastic leukaemias, uterus tumours, brain anaplastic oligodendromas	
CC	and uterus endometrial adenocarcinomas. The present sequence is the	

Aab41716 Human ORF
Aau28082 Novel hum
Abb57450 Human sec
Aab90656 Human sec
Abp41883 Human ova
Abb57451 Human sec
Aab90568 Human sec
Abg65430 Human alb
Abb71739 Drosophil
Abb67517 Drosophil
Aay97156 Aspergill
Aab40651 Human ORF
Add12509 PDZ ligan
Abb64827 Drosophil
Abb58814 Drosophil
Ade08492 Novel pro
Abg08365 Novel hum
Ade38413 Human pro
Aaw56699 Modified
Ade56061 Human Pro

26 274.5 7.3 868 3 AAB41716
27 274.5 7.3 868 4 AAU28082
28 270.5 7.2 735 5 ABB57450
29 268 7.1 921 4 AAB90656
30 268 7.1 921 5 ABP41883
31 264.5 7.0 735 5 ABB57451
32 264.5 7.0 868 4 AAB90568
33 264.5 7.0 868 5 ABG65430
34 262.5 7.0 1790 4 ABB71739
35 239 6.3 836 4 ABB67517
36 221 5.9 549 3 AAY97156
37 192 5.1 164 3 AAB40651
38 185 4.9 102 7 ADD12509
39 165.5 4.4 1970 4 ABB64827
40 165 4.4 1284 4 ABB58814
41 164 4.3 1736 7 ADE08492
42 156.5 4.1 129 4 ABG08365
43 156.5 4.1 942 7 ADE38413
44 155.5 4.1 942 2 AAW56699
45 155.5 4.1 942 7 ADE56061

CC protein of the invention. The gene is found on human chromosome 8

XX Sequence 723 AA;

Query Match 100.0%; Score 3774; DB 5; Length 723;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 723; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MILERPDGAGAGESPRLOISRRPKTRVSSLRGREGRLDYCAMRGCHVHGEPVR	60
Dy	1	MILERPDGAGAGESPRLOISRRPKTRVSSLRGREGRLDYCAMRGCHVHGEPVR	60
Qy	61	VHVGMPQLHAGVCDLSLTOCCQLOSRRAQIHQIDKELQMRGTGAENLYRATSNRVR	120
Dy	61	VHVGMPQLHAGVCDLSLTOCCQLOSRRAQIHQIDKELQMRGTGAENLYRATSNRVR	120
Qy	121	ETVALELSYNSNLQLLKEELELSGGVDPGRHSGEAVTVPMIPLGLKETKELDMSTPLK	180
Dy	121	ETVALELSYNSNLQLLKEELELSGGVDPGRHSGEAVTVPMIPLGLKETKELDMSTPLK	180
Qy	181	ELISVHFGEDGASVEAFIRELEALROAMRTSPRNESGLELTATYNNQICFLDARFLTPAR	240
Dy	181	ELISVHFGEDGASVEAFIRELEALROAMRTSPRNESGLELTATYNNQICFLDARFLTPAR	240
Qy	241	SLGLFFHWYDSLTGVPAAORALAFKGSVLNIGALHTQIGARQDRSCTEGARRAMEAFQ	300
Dy	241	SLGLFFHWYDSLTGVPAAORALAFKGSVLNIGALHTQIGARQDRSCTEGARRAMEAFQ	300
Qy	301	RAAGAFSLLRNFHSHAPSPDMSAASLCAIQLIMMAQAQECVPEGLSPASMAPQDCLAQL	360
Dy	301	RAAGAFSLLRNFHSHAPSPDMSAASLCAIQLIMMAQAQECVPEGLSPASMAPQDCLAQL	360
Qy	361	RLAQAAQAAEYRLVHRTMAQPPVHDYVPSVSWTALVHVKAEPFSLAHYHVAMALCDGS	420
Dy	361	RLAQAAQAAEYRLVHRTMAQPPVHDYVPSVSWTALVHVKAEPFSLAHYHVAMALCDGS	420
Qy	421	PATEGELPTHEQVFLQPTSSKPRGPVLPOELEERRQLGKAHLKRAILGQEPALRLHALC	480
Dy	421	PATEGELPTHEQVFLQPTSSKPRGPVLPOELEERRQLGKAHLKRAILGQEPALRLHALC	480
Qy	481	RVLREVDLLRAVISOTLQSLAKAYELDREDDFCBAEAPDIQPTKTHQKPEARMPLSQG	540
Dy	481	RVLREVDLLRAVISOTLQSLAKAYELDREDDFCBAEAPDIQPTKTHQKPEARMPLSQG	540
Qy	541	KGPDIFHRLGPLSVFSAKNRWELGVPVHLTRGEGGFLTLRGDSVPLIAAVIPGQAAAA	600
Dy	541	KGPDIFHRLGPLSVFSAKNRWELGVPVHLTRGEGGFLTLRGDSVPLIAAVIPGQAAAA	600
Qy	601	GLKEGDIYVSVNGQPCRWWRHAEVVTTELKAAEGASLQVWSLLPSSRLPSLGDRRPVLL	660
Dy	601	GLKEGDIYVSVNGQPCRWWRHAEVVTTELKAAEGASLQVWSLLPSSRLPSLGDRRPVLL	660
Qy	661	GPRLRLRSQREHGCKTPASTWASPRPLLNWSRKAQGGTGGCPQCAPVKPAPPSLKH	720
Dy	661	GPRLRLRSQREHGCKTPASTWASPRPLLNWSRKAQGGTGGCPQCAPVKPAPPSLKH	720
Qy	721	GWP 723	
Dy	721	GWP 723	

RESULT 2

ABB99400

ID ABB99400 standard; protein; 670 AA.

XX AC ABB99400;

XX AC ABB99400;

XX 29-JAN-2003 (first entry)

XX Amino acid sequence of a human intracellular signalling molecule.

XX Human; intracellular signalling molecule; INTS1G; Alzheimer's disease;

XX neurodegenerative disorder; Parkinson's disease; muscular disorder;

KW

myotonic dystrophy; catatonia; endocrine disorder; diabetes;
Grave's disease; cancer; leukemia; cervical cancer; breast cancers;
immunological disorder; scleroderma; systemic lupus erythematosus;
allergy; gastrointestinal disorder; Crohn's disease; renal disorder;
Goodpasture's syndrome; infection; cardiovascular disorder;
atherosclerosis; hepatic disease; cirrhosis.

Homo sapiens.

XX	Key	Location/Qualifiers
XX	Modified-site	16 /note= "potential phosphorylation site"
FT	Modified-site	35 /note= "potential phosphorylation site"
FT	Modified-site	52 /note= "potential phosphorylation site"
FT	Modified-site	58 /note= "potential phosphorylation site"
FT	Modified-site	159 /note= "potential phosphorylation site"
FT	Modified-site	161 /note= "potential phosphorylation site"
FT	Modified-site	163 /note= "potential phosphorylation site"
FT	Modified-site	234 /note= "potential phosphorylation site"
FT	Modified-site	259 /note= "potential phosphorylation site"
FT	Modified-site	370 /note= "potential phosphorylation site"
FT	Modified-site	387 /note= "potential phosphorylation site"
FT	Modified-site	503 /note= "potential phosphorylation site"
FT	Domain	513..592 /note= "PDZ domain"
FT	Modified-site	517 /note= "potential phosphorylation site"
FT	Modified-site	526 /note= "potential phosphorylation site"
FT	Region	528..530 /note= "cell attachment sequence"
FT	Modified-site	593 /note= "potential phosphorylation site"
FT	Modified-site	598 /note= "potential phosphorylation site"
FT	Modified-site	615 /note= "potential phosphorylation site"
FT	Modified-site	630 /note= "potential phosphorylation site"
FT	Modified-site	636 /note= "potential glycosylation site"
FT	Modified-site	638 /note= "potential phosphorylation site"
FT	Modified-site	663 /note= "potential phosphorylation site"

WO200277235-A2.

03-OCT-2002.

19-DEC-2001; 2001WO-US050315.

21-DEC-2000; 2000US-0257804P.

05-JAN-2001; 2001US-0260102P.

(INCY-) INCYTE GENOMICS INC.

Burford N, Ding L, Yue H, Thornton M, Walia NK, Gandhi AR;

Arvizu C, Baughn MR, Swarnakar A, Duggan BM, Lu DAW, Thangavelu K;

Warren BA, Tang YT, Khan FA, Yao MG, Emerling BM;

WPI; 2003-018929/01.


```
QY 61 VHVGMNPNQLHAGVCDLSLTQCCQLOSRRAQIHQIQIDKELQMTGAEINLYRATSNRVR 120
Db 27 -----KYGVSFVNQPCQLQSHRARLHQIQISKELRMTGAENLYRATSNWTWR 74

QY 121 ETVALELSYVNSNQLLKEELELSGGVDPGRHSGEAVTVPMIPLGLKETKELDWSPLK 180
Db 75 ETVALELSYVNSNQLLKEELELSGGVDPGRHSGEAVTVPMIPLGLKETKELDWSPLK 134

QY 181 ELISVHFGEDGASVEAFIRELEALRQAMRTSPRNSGLELLTAYYNOLCFIDARFLTPAR 240
Db 135 ELISVHFGEDGTSFETIQELEDLRQATRTSPRDEAGLDLAAVYSQCFIDARFFSPSR 194

QY 241 SLGLFFHWYDSLTGVPAAQORALAFKGSVLNFIQALHTQIGARQDRSCTEGARRAMEAFQ 300
Db 195 SPGLLFFHWYDSLTGVPAAQORALAFKGSVLNFIQALHTQIGARQDRSCTEGARRAMEAFQ 254

QY 301 RAAGAFSLRLNFSHAPSFDMSAASLCALEQLMMAQAQECVFEGLSPASMAPQDCLAOL 360
Db 255 RAAGAFSLRLNFSHAPSFDMSAASLCALEQLMMAQAQECVFEGLSPASMAPQDCLAOL 314

QY 361 RLAAQAAQVAAEYRLVHRTMAQPPVHDYVPVSWTALVHVKAIEYFSLAHYHVAALCDGS 420
Db 315 QLAQAAQVATEYGLVHRTMAQPPVHDYVPVSWTALVHVKAIEYFSLAHYHVAALCDGS 374

QY 421 PATEGELETHQVFLQPTSSKPRGPVLPOELEERRQLGKAHLKRAILQGEALRLHALC 480
Db 375 PA-KGELARQEHVF-QPSTPHEPLGFTLPQHPEDRRKLAKAHLKRAILQGEALRLHALC 432

QY 481 RVLREVDLLRAVISQTLORSIAKYAELREDDDPCFAAEAPDIQPKTHQKPEARMPRLSQG 540
Db 433 RVLKVDLLQVVTQALRRSLAKYSQLEREDDDFFATEAPDIQPKTHQTP 483

QY 541 KGPDIFHRLGPLSVFSKRNWRLVGPVHLTRGEGGGLTLRGDSVPLIAAVIPGSOAAAA 600
Db 484 -----GPLSVFSTKNRWQLVGPVHTRGEGGGLTLRGDSVPLIAAVIPGSOAAAA 534

QY 601 GLKEGDIYVSVNGQPCRWRRHAEVVTTELKAAAGEAGASLQVSWLLPSRLPSIGDRPVL 660
Db 535 GLKEGDIYVSVNGQPCRWRRHAEVVTTELKAAAGEAGASLQVSWLLPSRLPSIGDRPVL 594

QY 661 GPRGLLRQREHGCKTPASTWASRPPLNWSRKAQOGKGTGCPQRC 706
Db 595 W-----NQREGCFETPMTTRPWPILGWSRKNKQGTGSHDPDC 634

RESULT 4
ABM32685
ID ABM32685 standard; protein; 643 AA.
XX AC ABM32685;
XX DT 06-OCT-2003 (first entry)
XX DE Human Rhophilin amino acid sequence SEQ ID NO:27.
XX KW DGI-3; identification; cytosolic; gene therapy; cancer.
XX OS Homo sapiens.
XX FN WO2003033675-A2.
XX PD 24-APR-2003.
XX PF 18-OCT-2002; 2002WO-US033511.
XX PR 19-OCT-2001; 2001US-0342892P.
XX (DGTB-) DGI BIOTECHNOLOGIES INC.
XX PA Pillutla R, Blume AJ, Goldstein NI;
XX PI 2003-393525/37.
XX DR WPI; 2003-393525/37.
XX XX
```

```
PT New DGI-3 nucleic acid useful for preparing a composition for diagnosing
or treating cancer.
XX Example 6; Fig 4B; 215pp; English.
PS
CC The present invention describes an isolated nucleic acid (I) comprising:
CC (a) a sequence encoding A1, G3, D9, C5, D1, H2, H3, G12, E7 or A1 or
CC penetrating variants; (b) a sequence comprising at least 15 contiguous
CC nucleotides of (A); or (c) a sequence that is complementary or at least
CC 80% identical to (A) or (B). Also described: (1) a vector comprising the
CC nucleic acid; (2) a virus comprising the nucleic acid; (3) a host cell
CC comprising the vector; (4) a primer or probe comprising the nucleic acid;
CC (5) an isolated peptide comprising the amino acid sequence of A1, G3, D9,
CC C5, D1, H2, H3, G12, E7 or A1 or penetrating variants; (6) an antibody
CC that binds to the isolated peptide; (7) an isolated polypeptide complex
CC comprising the peptide and DGI-polypeptide complex; (8) a peptide library
CC comprising the peptide; (9) identifying a DGI-3 binder; (10) identifying
CC a DGI-3 antagonist agent; (11) diagnosing cancer; (12) a pharmaceutical
CC composition; (13) treating cancer; and (14) inhibiting activity in a
CC mammalian cell of a gene whose activity is regulated by DGI-3. (I) has
CC cytosolic activity and can be used in gene therapy. The nucleic acid (I)
CC can be used for preparing a composition for diagnosing or treating
CC cancer. ACF56944 to ACF57031 and ABM32656 to ABM32775 represent sequences
CC used in the exemplification of the present invention
XX
SQ Sequence 643 AA;
```

```
Query Match 61.8%; Score 2331; DB 6; Length 643;
Best Local Similarity 67.0%; Pred. No. 7e-213;
Matches 473; Conservative 49; Mismatches 112; Indels 72; Gaps 6;

QY 1 MILERPDCAGACGESPRLQISRRKPRKTRVSSURRGLRDVCAWRCGRVHGDDPVR 60
Db 1 MILERPDCQGTGESSRPQDD-----GSIR----- 26

QY 61 VHVGMNPNQLHAGVCDLSLTQCCQLOSRRAQIHQIQIDKELQMTGAEINLYRATSNRVR 120
Db 27 -----KYGVSFVNQPCQLQSHRARLHQIQISKELRMTGAENLYRATSNWTWR 74

QY 121 ETVALELSYVNSNQLLKEELELSGGVDPGRHSGEAVTVPMIPLGLKETKELDWSPLK 180
Db 75 ETVALELSYVNSNQLLKEELELSGGVDPGRHSGEAVTVPMIPLGLKETKELDWSPLK 134

QY 181 ELISVHFGEDGASVEAFIRELEALRQAMRTSPRNSGLELLTAYYNOLCFIDARFLTPAR 240
Db 135 ELISVHFGEDGTSFETIQELEDLRQATRTSPRDEAGLDLAAVYSQCFIDARFFSPSR 194

QY 241 SLGLFFHWYDSLTGVPAAQORALAFKGSVLNFIQALHTQIGARQDRSCTEGARRAMEAFQ 300
Db 195 SPGLLFFHWYDSLTGVPAAQORALAFKGSVLNFIQALHTQIGARQDRSCTEGARRAMEAFQ 254

QY 301 RAAGAFSLRLNFSHAPSFDMSAASLCALEQLMMAQAQECVFEGLSPASMAPQDCLAOL 360
Db 255 RAAGAFSLRLNFSHAPSFDMSAASLCALEQLMMAQAQECVFEGLSPASMAPQDCLAOL 314

QY 361 RLAAQAAQVAAEYRLVHRTMAQPPVHDYVPVSWTALVHVKAIEYFSLAHYHVAALCDGS 420
Db 315 QLAQAAQVATEYGLVHRTMAQPPVHDYVPVSWTALVHVKAIEYFSLAHYHVAALCDGS 374

QY 421 PATEGELETHQVFLQPTSSKPRGPVLPOELEERRQLGKAHLKRAILQGEALRLHALC 480
Db 375 PA-KGELARQEHVF-QPSTPHEPLGFTLPQHPEDRRKLAKAHLKRAILQGEALRLHALC 432

QY 481 RVLREVDLLRAVISQTLORSIAKYAELREDDDPCFAAEAPDIQPKTHQKPEARMPRLSQG 540
Db 433 RVLKVDLLQVVTQALRRSLAKYSQLEREDDDFFATEAPDIQPKTHQTP 483

QY 541 KGPDIFHRLGPLSVFSKRNWRLVGPVHLTRGEGGGLTLRGDSVPLIAAVIPGSOAAAA 600
Db 484 -----GPLSVFSTKNRWQLVGPVHTRGEGGGLTLRGDSVPLIAAVIPGSOAAAA 534

QY 601 GLKEGDIYVSVNGQPCRWRRHAEVVTTELKAAAGEAGASLQVSWLLPSRLPSIGDRPVL 660
Db 535 GLKEGDIYVSVNGQPCRWRRHAEVVTTELKAAAGEAGASLQVSWLLPSRLPSIGDRPVL 594
```

Db 535 GLKEDYIVSVNGQPCWKWKKLEVVTLQRLSMGEGYSLQVSVLLPSPERGTGPRRAALL 594
QY 661 GPRGLRSQREHGCKTPTASTWASPRLLNWSRKAQCKTGCGQPC 706
Db 595 W-----NORECGFETPMTTRTPWLPILGWSRKNQCKTGSHPDC 634
RESULT 5
ABR41761
ID ABR41761 standard; protein; 272 AA.
XX
AC ABR41761;
XX
DT 02-JUN-2003 (first entry)
XX
DE Human DTHP biochemical pathway protein.
XX
KW Human; dithp; diagnostic and therapeutic polynucleotide; diagnosis;
KW cancer; cell proliferative disorder; autoimmune disorder;
KW inflammatory disorder; infection; hormonal disorder; metabolic disorder;
KW neurological disorder; gastrointestinal disorder; transport disorder;
KW connective tissue disorder; drug screening; proteome analysis;
KW gene therapy; antisense therapy; genotyping; transgenic animal; knock in;
KW disease model; toxicological testing; transcript imaging;
KW biochemical pathway.
XX
OS Homo sapiens.
XX
FN WO200297031-A2.
XX
PD 05-DEC-2002.
XX
PF 27-MAR-2002; 2002WO-US010056.
XX
PR 28-MAR-2003; 2001US-0279619P.
PR 29-MAR-2003; 2001US-0280067P.
PR 29-MAR-2003; 2001US-0280068P.
PR 16-MAY-2003; 2001US-0291280P.
PR 17-MAY-2003; 2001US-0291829P.
PR 17-MAY-2003; 2001US-0291849P.
PR 19-JUN-2003; 2001US-0299428P.
PR 20-JUN-2003; 2001US-0299776P.
PR 20-JUN-2003; 2001US-0300001P.
XX
PA (INCYT) INCYTE GENOMICS INC.
XX
PI Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;
PI Dufour GE, Hillman JU, Yu JY, Tuason O, Yap PE, Anshey SR;
PI Daughtery SC, Dam TC, Liu TF, Nguyen DA, Kleefeld Y, Gerstin EH;
PI Peralta CH, David MH, Lewis SA, Chen AJ, Panzer SR, Harris B;
PI Flores V, Marwaha R, Lo A, Lan RY, Urashka ME;
XX
DR WPI; 2003-129518/12.
DR N-PSDB; ACC46698.
XX
PT Novel human diagnostic and therapeutic polypeptide useful for identifying
PT test compound which specifically binds to a polypeptide encoded by human
PT diagnostic and therapeutic polynucleotide, and to induce antibodies.
XX
PS Claim 27; SEQ ID NO 1296; 591pp; English.
XX
CC The invention relates to novel human diagnostic and therapeutic
CC polynucleotides designated dithp (ACC46080-ACC46749) and to their encoded
CC proteins (DITHP; ABR41136-ABR41812). The invention also relates to
CC polynucleotide sequences at least 90% identical to the dithp cDNA
CC sequences of the invention; recombinant vectors, host cells and
CC transgenic organisms comprising a dithp nucleic acid sequence; the
CC recombinant production of DITHP proteins; antibodies specific for DITHP
CC proteins; microarrays comprising dithp nucleic acid sequences; methods of
CC detecting dithp nucleotide and protein sequences; methods of screening
CC for compounds which specifically bind a DITHP protein; and methods of
CC assessing the toxicity of test compounds using a dithp hybridisation
CC probe. Dithp nucleic acid sequences and DITHP proteins may be used in the

CC diagnosis of a wide variety of conditions including cancer and other cell
CC proliferative disorders; autoimmune or inflammatory disorders; bacterial,
CC viral, fungal or parasitic infections; hormonal disorders; metabolic
CC disorders; neurological disorders; gastrointestinal disorders; transport
CC disorders; and connective tissue disorders. They may also be used to
CC screen for modulators of protein activity or gene expression. DITHP
CC proteins can additionally be used in analysis of the proteome of a tissue
CC or cell type and to induce antibodies. The dithp nucleic acids are
CC additionally useful in somatic or germline gene therapy of the disorders
CC mentioned above, as a source of antisense sequences, as a source of
CC probes and primers, in genotyping and identification of individuals, in
CC the generation of transgenic animal models of human disease or knock in
CC humanised animals, in toxicological testing, and in transcript imaging.
CC The present sequence represents a DITHP protein which is involved in a
CC biochemical pathway. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 272 AA;

Query Match 37.1%; Score 1401; DB 6; Length 272;
Best Local Similarity 100.0%; Pred. No. 1.2e-124;
Matches 265; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 459 GKAKLKRAILGQREALRLHALCRVLRVLDLRAVISQTLQRSIAKYAELDRDDFCFAAE 518
Db 8 GKAKLKRAILGQREALRLHALCRVLRVLDLRAVISQTLQRSIAKYAELDRDDFCFAAE 67
QY 519 APDIQPKTHQKPEARMPLRSQCKGPDIFHRLGPLSVFSAKNRWRLGVPVHLTRGEGFGL 578
Db 68 APDIQPKTHQKPEARMPLRSQCKGPDIFHRLGPLSVFSAKNRWRLGVPVHLTRGEGFGL 127
QY 579 TLRGDSPLVLIAAVPGSQAAAAGLKEGDYIVSVNGQPCRWHRHAEVVTTELKAGEAGASL 638
Db 128 TLRGDSPLVLIAAVPGSQAAAAGLKEGDYIVSVNGQPCRWHRHAEVVTTELKAGEAGASL 187
QY 639 QVVSLLPSSRLPSLGRDRPVLLGPRGLLSQREHGCKTPTASTWASPRLLNWSRKAQCKG 698
Db 188 QVVSLLPSSRLPSLGRDRPVLLGPRGLLSQREHGCKTPTASTWASPRLLNWSRKAQCKG 247
QY 699 TGGCQPCAPVVPKAPPSSLUKHPGW 723
Db 248 TGGCQPCAPVVPKAPPSSLUKHPGW 272

RESULT 6
AAB43571
ID AAB43571 standard; protein; 257 AA.
XX
AC AAB43571;
XX
DT 08-FEB-2001 (first entry)
XX
DE Human cancer associated protein sequence SEQ ID NO:1016.
XX

Human; cancer associated gene; cancer antigen; detection; cancer;
diagnosis; cytostatic; proliferative; vulnerable; immunomodulator;
antidiabetic; antiaschmatic; antirheumatic; antithrombotic; antiviral;
antiinflammatory; antithyroid; antiallergic; antibacterial; cardiac;
dermatological; neuroprotective; thrombolytic; coagulant; neutropenic;
vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;
immune disorder; haematopoietic cell disorder; autoimmune disorder;
allergic reaction; graft versus host disease; organ rejection;
haemostatic; thrombolytic; cardiovascular disorder; infection;
neurological disease; drug screening.

OS Homo sapiens.
XX
XX WO200055350-A1.
XX
XX 21-SEP-2000.
XX
XX 08-MAR-2000; 2000WO-US005882.
PF

```
XX 12-MAR-1999; 99US-0124270P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Ruben SM;
XX WPI; 2000-587533/55.
XX N-PSDB; AAC77780.
XX Novel isolated nucleic acids comprising sequences encoding peptides
XX useful for treating or diagnosing e.g. cancer.
XX Claim 11; Page 1597-1598; 2352pp; English.
XX
XX AAC77607 to AAC78448 encode the human cancer associated proteins given in
XX AAB43398 to AAB44239. The proteins can have activities based on the
XX tissues and cells the genes are expressed in. Example of activities
XX include: cytostatic; proliferative; vulnerary; immunomodulator;
XX antidiabetic; antiaesthetic; antirheumatic; antiarthritic;
XX antiinflammatory; antithyroid; antiallergic; antibacterial; antiviral;
XX dermatological; neuroprotective; cardiatic; thrombolytic; coagulant;
XX nototropic; vasotropic; antipsoriatic and antiangiogenic. The
XX polynucleotides and polypeptides can be used for preventing, treating or
XX ameliorating medical conditions and diagnosing pathological conditions.
XX polynucleotides, polypeptides, antibodies, agonists and antagonists from
XX the present invention may be used to treat immune disorders by activating
XX or inhibiting the proliferation, differentiation or mobilisation of
XX immune cells; to treat disorders of haematopoietic cells, autoimmune
XX disorders, allergic reactions, graft versus host disease and organ
XX rejection, modulate haemostatic or thrombolytic activity, modulate
XX inflammation, cancers, cardiovascular disorders, neurological disease and
XX bacterial or viral infections. The peptides, nucleotides, antibodies,
XX agonists and antagonists may be also be used in drug screens. AAC78449 to
XX AAC78457 and AAB44240 represent sequences used in the exemplification of
XX the present invention
XX Sequence 257 AA;
XX
XX Query Match 33.9%; Score 1278; DB 3; Length 257;
XX Best Local Similarity 99.2%; Pred. No. 5.9e-113;
XX Matches 245; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
QY 459 GKXHLKRALGQEEALRHLCRVLRVLDLRAVISQTLQRLSLAKYAEIRDREDDFCEAAE 518
Db 8 GKXHLKRALGQEEALRHLCRVLRVLDLRAVISQTLQRLSLAKYAEIRDREDDFCEAAE 67
QY 519 APDIQPKTHQKPEARMPLRSQKGGPDI FHLGLPLSVFSKKNRWLVGPVHLTRGEGGGL 578
Db 68 APDIQPKTHQKPEARMPLRSQKGGPDI FHLGLPLSVFSKKNRWLVGPVHLTRGEGGGL 127
QY 579 TLRGDSPVLIAAVIPGQAAAAAGLKEGDYIVSVNGQPCRWHRHAEVVTLEKAAAGAGASL 638
Db 128 TLRGDSPVLIAAVIPGQAAAAAGLKEGDYIVSVNGQPCRWHRHAEVVTLEKAAAGAGASL 187
QY 639 QVVSLLPSRLPSLGRDPRVLLGPRLLRSQREHGCKTPTASTWASPRLLNWSRKAQQCK 698
Db 188 QVVSLLPSRLPSLGRDPRVLLGPRLLRSQREHGCKTPTASTWASPRLLNWSRKAQQCK 247
QY 699 TGGCFQP 705
Db 248 TGGCFSP 254
XX
RESULT 7
AAU00869
ID AAU00869 standard; protein; 257 AA.
XX AAU00869;
XX
XX 04-JUL-2001 (first entry)
XX Human cancer related protein 5.
XX
```

```
XX Human; cancer related protein; food additive; preservative; immunogen;
KW antibody; bone cancer; adrenal cancer; bone marrow cancer; breast cancer;
KW gastrointestinal cancer; liver cancer; lung cancer; urogenital cancer;
KW immune disorder; Addison's disease; allergy; diabetes mellitus;
KW autoimmune haemolytic anaemia; autoimmune thyroiditis; Crohn's disease;
KW multiple sclerosis; rheumatoid arthritis; ulcerative colitis;
KW acquired immunodeficiency syndrome; AIDS; cardiovascular disorder;
KW myocardial ischaemia; wound healing; neurological disorder;
KW Parkinson's disease; Alzheimer's disease; cerebral anoxia; epilepsy;
KW viral infection; bacterial infection; fungal infection;
KW parasitic infection; agonist; antagonist.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Region 69..81
XX /label= Immunogenic_epitope
XX Region 84..91
XX /label= Immunogenic_epitope
XX Region 106..112
XX /label= Immunogenic_epitope
XX Region 216..224
XX /label= Immunogenic_epitope
XX Region 239..250
XX /label= Immunogenic_epitope
XX WO200118014-A1.
XX 15-MAR-2001.
XX 30-AUG-2000; 2000WO-US023794.
XX 03-SEP-1999; 99US-0152296P.
XX 06-OCT-1999; 99US-0158003P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Roschke V;
XX WPI; 2001-235186/24.
XX N-PSDB; AAS00831.
XX Twenty nine nucleic acid molecules encoding human cancer associated
XX proteins, useful in the prevention, treatment and diagnosis of cancer,
XX immune disorders, cardiovascular disorders and neurological diseases.
XX Disclosure; Page 394-395; 427pp; English.
XX
XX The sequence represents a novel Human cancer related protein. The
XX polynucleotides and polypeptides are useful for preventing, treating or
XX ameliorating a medical condition in e.g. humans, mice, rabbits, goats,
XX horses, cats, dogs, chickens or sheep. The polypeptides can also be used
XX as a food additive or preservative to increase or decrease storage
XX capabilities. The polynucleotide are useful for chromosome
XX identification. The nucleic acids, protein, antibodies, agonists and
XX antagonists are useful in the diagnosis, treatment and prevention of
XX cancer (e.g. cancers of the adrenal gland, bone, bone marrow, breast,
XX gastrointestinal tract, liver, lung, or urogenital), immune disorders
XX (e.g. Addison's disease, allergies, autoimmune haemolytic anaemia,
XX autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple
XX sclerosis, rheumatoid arthritis and ulcerative colitis, acquired
XX immunodeficiency syndrome, AIDS), cardiovascular disorders such as
XX myocardial ischaemias, wound healing, neurological diseases (e.g.
XX Parkinson's disease, Alzheimer's disease, cerebral anoxia and epilepsy)
XX and infectious diseases such as viral, bacterial, fungal and parasitic
XX infections. Numerous examples of each type of disorder are given in the
XX specification
XX Sequence 257 AA;
XX
XX Query Match 33.9%; Score 1278; DB 4; Length 257;
XX Best Local Similarity 99.2%; Pred. No. 5.9e-113;
```


DT 02-DEC-2002 (first entry)
XX Human GTP-Rho binding protein 2.
XX
KW Human; GTP-Rho binding protein 2; GRBP2; chromosome 19q12; antigen;
KW oncogene; tumour; liposarcoma; ichthyosis congenita III; vaccine;
KW benign familial infantile convulsion; gene therapy.
XX
OS Homo sapiens.
XX
XX EP1231216-A2.
XX
XX 14-AUG-2002.
XX
XX 17-JAN-2002; 2002EP-00001026.
XX
XX 30-JAN-2001; 2001WO-US000663.
XX 30-JAN-2001; 2001WO-US000664.
XX 30-JAN-2001; 2001WO-US000665.
XX 30-JAN-2001; 2001WO-US000666.
XX 30-JAN-2001; 2001WO-US000667.
XX 30-JAN-2001; 2001WO-US000668.
XX 30-JAN-2001; 2001WO-US000669.
XX 30-JAN-2001; 2001WO-US000670.
XX 29-JUN-2001; 2001US-00895040.
XX (AECOM-) AECOMICA INC.
XX
XX Shannon ME, JI Y;
XX WPI; 2002-684026/74.
XX P-PSDB; ABS71855, ABS71856.
XX
XX Novel GTP-Rho binding protein 2 and nucleic acids encoding the protein,
XX useful for the manufacture of a medicament for treating a disease
XX associated with altered expression or activity of human GRBP2 protein.
XX
XX Claim 10; Fig 3; 101pp; English.
XX
XX The invention relates to an isolated GTP-Rho binding protein 2 (GRBP2)
XX polypeptide or a fragment of at least 6 amino acids or a sequence in
XX which at least 95% of deviations from GRBP2 sequences are conservative
XX substitutions. Also included are an isolated nucleic acid (GRBP2 NA)
XX encoding GRBP2 comprising the full length cDNA or CDS, fragments or
XX variants, GRBP2 vectors, host cells, antibodies, transgenic non-human
XX animals modified to contain GRBP2 NA (or unable to express the endogenous
XX orthologue of GRBP2), diagnosing a disease caused by a mutation in human
XX GRBP2 or altered expression of GRBP2, ant-agonists of GRBP2, GRBP2
XX microarrays, fusion proteins and screening for agents that modulate the
XX expression of GRBP2 NA. GRBP2 is useful for identifying binding partners
XX of GRBP2. GRBP2, GRBP2 NA and Ab are useful in therapy and in the
XX manufacture of a medicament for the treatment or prevention of a disorder
XX associated with increased or decreased expression or activity of human
XX GRBP2 (e.g. tumours, liposarcoma, ichthyosis congenita III and benign
XX familial infantile convulsion, all associated with the chromosomal
XX location of GRBP2, 19q12). GRBP2 is useful as a standard in immunoassay
XX specific for the proteins, to be used in a therapeutic agent, as
XX vaccines, to be and as antigens (e.g. for epitope mapping) or immunogens
XX (e.g. for raising antibodies). GRBP2 NA is useful as hybridisation probes,
XX to prime synthesis of nucleic acids, to prime first strand cDNA sequences,
XX on an mRNA template, and to drive in vivo expression of the proteins. The
XX vector is useful for shuttling GRBP2 NA between host cells derived from
XX disparate organisms, for inserting GRBP2 NA into host cell chromosome,
XX for expressing sense or antisense RNA transcripts of GRBP2 NA in vitro or
XX within a host cell, and for expressing GRBP2 alone or as fusions to
XX heterologous polypeptides. The antibody is useful as an analytical
XX reagent for detection and quantification of GRBP2 and as an immuno
XX therapeutic agent and is useful for flow cytometric detection, for
XX scanning laser cytometric detection, or for fluorescent immunoassay. The
XX present sequence is a GRBP2 protein sequence
XX Sequence 686 AA;

Query Match 33.1%; Score 1249; DB 5; Length 686;
Best Local Similarity 45.8%; Pred. No. 1-7e-109;
Matches 265; Conservative 107; Mismatches 197; Indels 10; Gaps 5;
QY 74 GDSLTQIQGQLOSPRAQHQQIDKELQMRCAENLYRATSNRRVRETVALBELSVNSN 133
DB 24 GGNPLAQTGRSLQONQRAALNQOILKAVRMIRGAENLLKVATNSKVRREQRLBELSFVNSD 83
QY 134 LQLKEELEELSGGVDPRGRHGEAVTPMIPGLKTKELDMSTPLKELISVHFGEDEGAS 193
DB 84 LQMLKEELEGLNISVGVIQNTTEAFTIPLIPGLKETKDVDFAVVLKDFILEHYSDEGYL 143
QY 194 YEABIREALRQAMTPSPRNSGLELLTAYYNQLCFLDARFUTPARSIFGLFFHMYDSLT 253
DB 144 YEDEIADLMDLRQACRTPSRDEAGVELLMTYFTQLGFVESRFPPTQMGILLFTWYDSLT 203
QY 254 GYPAQORALAFKESVLENIIGALHTOIGARQDRSCTEGARRAMEAFORAGAFSLIRENF 313
DB 204 GYFVSQONLLLEKASVLENTGALYITQIGTRCDRQOAGLESALDAFORAAGVINYLKDTF 263
QY 314 SHAPSDMSAASICALLEQLMMAQAQCEVFEGLSPSPASMAPQDCLAQLRLAQEAQVAAEV 373
DB 264 THTPSYDMSPAMLSVLVKMMLAQAESVFEKISLPG--IRNEFFMLVKVQAQEAQVAAEV 321
QY 374 RLVHRTMAQPFVHDYVPSWTALVHVKAIEYFRSLAHYHVAMALCDG--SPATEGELPTHE 431
DB 322 QQLHAAMSQAQPVKENIPYSWASLACVKAHYAALAHYFTAILIIDHQVKPQGTDLD---HQ 378
QY 432 QVFLQPTTSSKPRG--PVLP-OELERRQLGKAHLKRAILGQFEALRLHALCVLREVDL 488
DB 379 EXKLSQLYDHMEGLTPLATLKNQORQLGKSHLRRVAHHEESVREASLCKKLSIEV 438
QY 489 LRRAVISQTLQRSKAYAELEDDREDDFCEAAEAPDIQKTHQKPEARMPLRSQKGPDIPIHR 548
DB 439 LQKVLCAAQERSRLTYAQHQEEDDLNLIDASVVAKTEQEVNDIILPQFSKLTVTDTFFQK 498
QY 549 LGPLSVFSKKNRWLVGPHVLTGRGGFGLTRGSGSPVLIAAVIPGSOAAAAGLKEGDYI 608
DB 499 LGPLSVFSKANKWTTPRSIRFTAEEGDLGFTLRGNAPQVHFLLDPYCSASVAGAREGYI 558
QY 609 VSVWGQPCRWHRHAEVVTELLKAAGAGASLQVVSLPSS 647
DB 559 VSIQLVDCXKWLTLSEVMKLLKSFGEDEIEMKVVSLLDST 597
RESULT 11
ID AAB68561
XX AAB68561 standard; protein; 666 AA.
AC AAB68561;
XX
DT 19-APR-2001 (first entry)
XX Human GTP-binding associated protein #61.
DE
XX
KW Human; guanosine triphosphate binding associated protein; GTP; GRAP;
KW inflammation; AIDS; Addison's disease; anaemia; arteriosclerosis; asthma;
KW osteoporosis; psoriasis.
XX
OS Homo sapiens.
XX WO200105970-A2.
XX
XX 25-JAN-2001.
XX
XX 19-JUL-2000; 2000WO-US019698.
XX
XX 19-JUL-1999; 99US-0144595P.
XX 23-AUG-1999; 99US-0150460P.
XX 15-OCT-1999; 99US-0159849P.
XX
XX (INCY-) INCYTE GENOMICS INC.

PR 14-NOV-2001; 2001US-0333184P.
PR 14-NOV-2001; 2001US-0333272P.
PR 21-NOV-2001; 2001US-0332694P.
PR 03-DEC-2001; 2001US-0337426P.
PR 04-DEC-2001; 2001US-0338092P.
PR 04-DEC-2001; 2001US-0337185P.
PR 03-JAN-2002; 2002US-0345705P.
PR 07-MAR-2002; 2002US-00092900.
XX (CURA-) CURAGEN CORP.
XX
PI Padigaru M, Spytek KA, Shenoy SG, Taupier RJ, Pena CEA, Li L;
PI Zerhusen BD, Gusev V, Ji W, Gorman L, Miller CE, Kekuda R;
PI Patturajan M, Gangolli E, Vernet CAM, Guo X, Tchernev V;
PI Fernandes ER, Casman SJ, Malyankar UM, Gerlach V, Liu Y, Anderson D;
PI Spaderna SK, Catterton E, Burgess C, Leite M, Zhong H, Alsobrook JP;
PI Lepley DM, Rieger DK;
XX WPI; 2002-723332/78.
DR N-PSDB; ABX97192.
XX
XX NOVX polypeptides and polynucleotides, useful for preventing or treating
PT a disorder associated with aberrant NOVX expression or activity e.g.,
PT cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial
PT asthma.
XX
XX Claim 1; Page 470-471; 1103pp; English.
XX
CC This invention describes novel human NOVX polypeptides which have
CC cytostatic, cardiant, antiarteriosclerotic, antiasthmatic and hypotensive
CC activity. Pharmaceutical compositions comprising the NOVX proteins or
CC nucleic acid molecules or NOVX antibodies are useful for preventing or
CC treating a disorder associated with aberrant NOVX expression or activity
CC e.g. cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial
CC asthma. The products of the invention can be used for gene therapy or in
CC a vaccine. AB065041-AB065218 represent the NOVX polypeptides encoded by
CC ABX97008-ABX97185
XX
SQ Sequence 686 AA;

Query Match
Best Local Similarity 31.7%; Score 1195; DB 5; Length 686;
Matches 263; Conservative 104; Mismatches 196; Indels 22; Gaps 8;

QY 74 GCDSTLTQCGQSPRAHQHQIDKELQMTGCAENLYRATSNRVRETVALELSVNSN 133
DB 24 GGNPLAQTRSKLQORALNQOILKAVEMRTGAENLLKVATNSKVREQVRLESPNSD 83
QY 134 LQLKEELEELSGGVDPRGHGSEAVTPMPLGLKETKELDSTPLKELISVHFEGDGAS 193
DB 84 LQMLKEELEGLNISGVVQNTTEATIPPLGLKETKDQDFAVVLKDFILEHSEGGYL 143
QY 194 YBAETRELEALQAMRTPSNBSGLELLTPAYYNQLCFDLARELTTPARSIGLFFHWDSLT 253
DB 144 YEDELADLMDLQACRTSRDEAGVELLMTYFIQLGFVSRFPPTQWGLLFTWDSLT 203
QY 254 GYPAQORALAFKGSVLENGALHTQIGAROBSCTEGARRAMEAPQRAAGFSLRENF 313
DB 204 GFVFSQNNLLKASVLFNTGALYTIQIGTRCDRTQAGLESIDAIPQRAAGVLYNKDTF 263
QY 314 SHAPSPDMSAALCALQELMLQAQECVPEGLSPSPASMAPQCLQALRLAQEAQVAAAY 373
DB 264 THTPSYDMSPAMLSVLVKWMLAQAESVEKISLPG--IRNFFMLVKVQAQAAGVEVY 321
QY 374 RLIVHTMAQPPVHDVYVPSWTLVHVKAEIFRSLAHYHVAMALCDG--SPATEGELPTE 431
DB 322 QQLHAAMSQAQKIPYNSWASLACVKAHYAALAHYFTAILLDHVKVPGTDLD--HQ 378
QY 432 QVFLQPTTSKPRG--FVLPL-QELERRQLGKHLKRALGQBEALRLHALCRVREVDL 488
DB 379 ECKLSQLYDHPGELTFLATLKNQDQRRQLGKSHLRRAMAHEESVREASLCKLRSIEV 438
QY 489 LRAVISQTLQRLAKYAEILDREDDFCEAAEAPDI-----QPKTHQKPEARPLSQGKG 542

DB 439 LQKVLCAAQERSRLTYAQHQEEDLNLIDAPRVILLKSLRTLXCPSS--PADSHGLL 496
QY 543 POIFHRIGPLSVFSKAKNRWRLVGPVHLTRGEGGFGTLTRGDSPLVLTAAVIFGSCAAAAGL 602
DB 497 FE----AGPLSVLSANKRWTPRPSIRFTAEGDLGFTLRGNAEPVQHFLDYPYCSAVAGA 552
QY 603 KEGDYIVSYNGQPCRWRRHAEVVVELKAAAGEAGASLQVVSLLPSS 647
DB 553 REGDYIVSIQLVDCKWLTLSVWMLKLSFGDEIEKMKVSLDST 597
RESULT 13
ABG08369
ID ABG08369 standard; protein; 1111 AA.
XX
AC ABG08369;
DT 13-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #8360.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS72556.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 20; SEQ ID NO 38728; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1111 AA;

```
Query Match      26.7%; Score 1007; DB 4; Length 1111;
Best Local Similarity 39.7%; Pred. No. 4.9e-86;
Matches 238; Conservative 99; Mismatches 212; Indels 50; Gaps 12;

QY 110 LYRATSNRRVETVALELSYVNSLQLKEBLEELSGGVDFGRHGEAVTVPMIPLGLKE 169
DB 23 LLRYAANPKVEQVRLELSFVNSDLQMLKELEGLN-----ISGVYQ 65

QY 170 TKELDWSTPLKELISVHFGEGDGYAEIRLEALRQAMRTPSPNESGLELLTAYYNQLC 229
DB 66 NTD-----EDGYLYDEITDMLRQACRTPESRDEARVELLMYFIQLG 109

QY 230 FLDAFLTPARSGLGFHHWYDSLTGVPAQALAFKGSVLFNFCALHTQICARODRST 289
DB 110 FVENRFPPTWQMGFLFTWYDSLTGVLSQQNLLLEKASVLFNTGVLTQIGTRYRHTQ 169

QY 290 EGARRAMEAFQRAAGAFSLLENFSPMSAASLCALQMLMAQORCVFEGLSPPA 349
DB 170 AGLQSAIDAFQRAAGVLYNKETFTFHPNSYDMIPAMLSVLVKMMLAQTESVFEKISLPG 229

QY 350 SMAQDCLAQLRLAQRAAQVAEYELVHRTWAQPPVHDYVPSVTALVHVKAEPERSLAH 409
DB 230 IW--NEFPMLVKVQAEAAKVGVEYQQQLHQMSQAPVKENIPYSWASLACVKAHYAALAH 287

QY 410 YHVAMALCDG--SPATEGELPHEQVFLQPPTSSKPRG--FVLP--QEELEERRQLGKAHLK 464
DB 288 YFTAILLLDHQVKPTGLD---HQEKLSQLYDHMPGLTPLATLKNQOQRRLGKSHLR 344

QY 465 FALQGEALRLHALCRVLREVDLLRAVISQTLORSIAKYAFLEDRDPFCRAEAAPDLP 524
DB 345 RAMAHHEBSVREASLCKKLRSLEVLQKVLCAQERSRLTYAQHQBEDDLLMLIDAPSVA 404

QY 525 KTHQPEARMPLSQGKGPDIHRLGLPLSVFSAKNRWRLVGPVHLTRGEGFGFLTRGDS 584
DB 405 KTEQEVDIILPQFSKLTVTDFQKLGPLSVFSANKRWTTPPSIRFTAEEDGLGFTLRGNA 464

QY 585 PVLIAAVIPGSAAGLKEGDIYIVSVNGQPCQRRWRHAEVVTTELKAAGEAGASLOWSL 644
DB 465 PQVHFPLDPCYASVAGAREGDIYIVSIQLVDCKWLTLSEVMKLLKSFGEDIEMKVVSLL 524

QY 645 PSSRLPSLGRDPRVILGPRGLRSOREHGCKTPTASTWASPRLLNWSKAQOQKGGCP 703
DB 525 DSTH--PCLA--RGFGLRVXAVL--NRTASSPMPPTQLRPRP--NW--QSEMGRRSSLP 576

RESULT 14
ABG61936
ID ABG61936 standard; protein; 556 AA.
AC
AC ABG61936;
DT 15-AUG-2002 (first entry)
XX
XX Prostate cancer-associated protein #137.
DE Prostate cancer; prostate tumour tissue; human; mammal; cytostatic.
KW
XX
XX Mammalia.
XX
XX WO20020268-A2.
PD
XX
XX 18-APR-2002.
XX
XX 12-OCT-2001; 2001WO-US032045.
XX
XX 13-OCT-2000; 2000US-00687576.
PR 08-DEC-2000; 2000US-00733288.
PR 08-DEC-2000; 2000US-00733742.
PR 24-JAN-2001; 2001US-0263957P.
PR 16-MAR-2001; 2001US-0276791P.
PR 16-MAR-2001; 2001US-0276888P.
PR 06-APR-2001; 2001US-0281922P.

24-APR-2001; 2001US-0286214P.
30-APR-2001; 2001US-00847046.
04-MAY-2001; 2001US-0288589P.
(EOSB-) EOS BIOTECHNOLOGY INC.
Gish KC, Mack DH, Wilson KE, Afar D, Hevezi P;
WPI; 2002-471335/50.
N-PSDB; ABK92255.
Detecting a prostate cancer-associated transcript in a cell in a patient,
useful for diagnosing prostate cancer (PC) or screening modulators of PC,
by determining if prostate cancer-associated genes are expressed in a
prostate tissue.
Claim 27; Page 418; 436pp; English.
The present invention relates to methods of detecting a prostate cancer-
associated transcript in a cell from a patient. The method comprises
contacting a biological sample from the patient with prostate cancer-
associated polynucleotides (designated PC genes) that selectively
hybridise to a sequence that is at least 80% identical to them. The
prostate cancer-associated polynucleotide sequences are differentially
expressed in prostate tumour tissue or in prostate cancer and are derived
from the tissues of various organisms such as humans or other mammals
(e.g. mice, sheep and dogs). The methods of the invention are useful for
diagnosing and treating prostate cancer in mammals. The prostate cancer-
associated genes are useful for diagnosing or treating prostate cancer,
as well as for identifying modulators of prostate cancer or agents that
inhibit prostate cancer. The nucleic acid sequences are particularly
useful in gene therapy, as a vaccine or in antisense applications.
ABG61800-ABG61944 represent prostate cancer-associated proteins
XX
SQ Sequence 556 AA;
Query Match      25.8%; Score 972; DB 5; Length 556;
Best Local Similarity 44.1%; Pred. No. 3.5e-83;
Matches 208; Conservative 83; Mismatches 171; Indels 10; Gaps 5;

QY 181 ELISVHFEDGASVBAEIRELEALRQAMRTPSPNESGLELLTAYYNQLCFLDARLTPAR 240
DB 1 DFLEHYSEDGVLVEDEIADLMDLQACRTPSRDEAGVELLMYFIQLGFVESRFFPPT 60

QY 241 SIGLFFHWYDSLTGVPAQALAFKGSVLFNFCALHTQICARQDRSTEGARRAMEAFQ 300
DB 61 QMGLLFTWYDSLTGVPSQQNLLLEKASVLFNTGALYTIQIGTRCDRQTOAGLESAIDAFQ 120

QY 301 RAAGAFSLLENFSPMSAASLCALQMLMAQOEVCVEGSPASMAPOCLAQL 360
DB 121 RAAGVLYNLKDTFTHTPSYDMSPAWLSVLVKNMLAQAESVFEKISLPG--IRNEFFMLV 178

QY 361 RLAEQAAQVAEYELVHRTWAQPPVHDYVPSVTALVHVKAEPERSLAHYHVMALCDG- 419
DB 179 KVAQEAQKVGVEYQQHAAMSQAPVKENIPYSWASLACVKAHYAALAHYFTAILLDHQ 238

QY 420 -SPATEGELPHEQVFLQPPTSSKPRG--FVLP--QEELEERRQLGKAHLKRAILGOBEALR 475
DB 239 VKPGTGLD---HQEKLSQLYDHMPGLTPLATLKNQOQRRLGKSHLRAMAHHEBSVR 295

QY 476 LHALCRVLREVDLLRAVISQTLORSIAKYAELEDRDDFCEAAEAPDIQPKTHQKPEARM 535
DB 296 EASLCKKURSIEVLQKVLCAQERSRLTYAQHQBEDDLLMLIDAPSVAKTQEVDIILP 355

QY 536 RLSSQKGPDIHRLGLPLSVFSAKNRWRLVGPVHLTRGEGFGFLTRGSPVLIIAVIPS 595
DB 356 QFSKLTVTDFQKLGPLSVFSANKRWTTPPSIRFTAEEDGLGFTLRGNAPVQVHFLDPYC 415

QY 596 QAAAGLKEGDIYIVSVNGQPCQRRWRHAEVVTTELKAAGEAGASLOWVSLPSS 647
DB 416 SASVAGAREGDIYIVSIQLVDCKWLTLSEVMKLLKSFGEDIEMKVVSLLDST 467
```

RESULT 15

AB63879
ID ABB63879 standard; protein; 648 AA.
XX AC ABB63879;
XX 26-MAR-2002 (first entry)
XX Drosophila melanogaster polypeptide SEQ ID NO 18429.
DE Drosophila melanogaster polypeptide; cell signalling; insecticide;
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX Drosophila melanogaster.
OS Drosophila melanogaster.
XX WO200171042-A2.
XX 27-SEP-2001.
XX 23-MAR-2001; 2001WO-US09231.
XX 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX (PEXE) PE CORP NY.
XX Venter JC, Adams M, Li PWD, Myers EW;
PI WPI; 2001-656860/75.
DR N-PSDB; ABL07982.
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions.
XX Disclosure; SEQ ID NO 18429; 2ipp + Sequence Listing; English.
PS The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 648 AA;
Query Match 21.6%; Score 817; DB 4; Length 648;
Best Local Similarity 31.7%; Pred. NO. 2.9e-68;
Matches 215; Conservative 114; Mismatches 233; Indels 116; Gaps 19;
QY 98 DKELQRTGAENLYRATSNRRVRETVL-ELSY- ---VNSNLQLKLEELSGGVDPGR 152
DB 3 EKEHELST- ---RVVKNNEEDDEKLNELSCFVRINSSVEIYQSE- --- 44
QY 153 HGEAVTVPMIPGLKTKELDWSTPLKELISVHFGEDGASYEAEIRELEALRQAMRTPS 212
DB 45 --SHNGIMPMIPGLKTKELINFMPEPDSFILEHVSSEPSMYIDAIDMTDTRQASKTPS 102
QY 213 RNESGLELITAYNQICFLDARFLTPARSGLGFPHWYDSLTGVPQAORALAFKGSVLN 272
DB 103 RDALGVALLFRYNTLYYVERFFPPDRNLGVYFEWYDSLTGVPSCQRTIAFEKACTLEN 162
QY 273 IGAHTQIGARODRSCTEGARAMEAFORAAGAFSLLENFSHAPSPDMSAASLCALEOL 332
DB 163 LGGIYTIQIGARDHRTTTERGLDLAVDSFLRAAGVFRHIYDTFTNAPSMOLKPPQVLDVLSL 222
QY 333 MMAQAQECVFEGLSPSPASMAPQDCLAQLRAQAAQAAVAAEYRLVHRTMAQPPVHDYVPVS 392
DB 223 MLSQARECLFEKLQLQIEAMSHDCQAFRDLAGAAQISHEYNEMHKNIQANDHTHYLPEC 282

QY 393 WTALVHVKAERYFRSLAHYHVAMALCDGSPATEGELPHEQVFLQPPTTSKPRGEVLPQEL 452
DB 283 WAGLVFVKAELYKAFAYFYKARS- ---DATD-ELKASK- ---SSQKNQSSFIGNSQ 330
QY 453 EERR- ---QLGK- ---AHLKRAILGOBEALRIHALCRVIRE- ---VDLLRAV 492
DB 331 EVERITTADYGASDEASTSIANKLAHLKEALASIEEAQELQRMCRFLKNKASLTYMKEV 390
QY 493 ISQTLQRSIAKY- ---AELDREDDFCEAAEAPDIQPKTHQKPEARMPLRSQCKGPDIFH 547
DB 391 HSKS-QEELEKFRLOQASAKNIEDGDLERS- ---VEASSKFTLSLTGPDFTSHKVDPEK 445
QY 548 RLGPLSVFSAKNRWRLVGPVHLTRG- ---EGG- --- 575
DB 446 RLGPFAIFSARRHWTAPRCVRLQKGSLSYHVSNDNKCPLDNDDDEHDGGVLYKEEF 505
QY 576 --FGLTLRGDSPVLIAAIVIPSGQAAAAGLKEGDYIVSVNGQPCRWHRHAEVVTTELKAAGE 633
DB 506 ENFGFHVRGDAPVITAHVEINSLADLGGIKEGDFIVEIAGVDVKWYSHQQVQLIQSCGS 565
QY 634 AGASLQVWSLLPSSRLPSLGDRRPVLLGPRGLLRSQREHGCKT- ---PASTWASPRPLLNWS 691
DB 566 T-LELRVITPMDRNYLKPILSSK- ---GSLSTLSAASSSGISSGFPSPSTIAAKPKLHL- 618
QY 692 RKAQOKTGGCPCPCAPV 709
DB 619 -----KTSSSRPAGSV 630

Search completed: July 13, 2004, 12:04:25
Job time : 91 secs

mis Page Blank (uspto)

QY	121	CTGAGAGAGTGTGTGCTGGAGAGGTTCTGGGTCCA	CAGAGGGGAAGACCCAGTGCCT	180	QY	1201	GGCGAGTACTTCGGTCCCTGGCCCACTACCACTAGCAT	GGCCCTCTGGCAGCGCTCC	1260
Db	223	CTGAGAGAGTGTGTGCTGGAGAGGTTCTGGGTCCA	CAGAGGGGAAGACCCAGTGCCT	282	Db	1303	GGCGAGTACTTCGGTCCCTGGCCCACTACCACTAGCAT	GGCCCTCTGGCAGCGCTCC	1362
QY	181	GTGACGTTGGCCCCCATGATCGGACGCTTCATGCA	TGGGCTGTCACTCCCTGACGCAG	240	QY	1261	CCAGGACCCGAGGGAGAGCTCCCGACGACGACGAG	GTCTTCTTGACAGCCCCCAGCTCC	1320
Db	283	GTGACGTTGGCCCCCATGATCGGACGCTTCATGCA	TGGGCTGTGACTCCCTTGACGCAG	342	Db	1363	CCAGGACCCGAGGGAGAGCTCCCGACGACGACGAG	GTCTTCTTGACAGCCCCCAGCTCC	1422
QY	241	ATCCAGTGGCGCCAGCTGCAGAGCCGACAGGCCAG	ATTCACAGCAGATTGCAAGGAG	300	QY	1321	TCTAGCCCCGAGGCCCTGTGTCGCGCAGGAGCTG	AGGCGCAGGAGCTTGGCAAG	1380
Db	343	ATCCAGTGGCGCCAGCTGCAGAGCCGACAGGCCAG	ATTCACAGCAGATTGCAAGGAG	402	Db	1423	TCTAGCCCCGAGGCCCTGTGTCGCGCAGGAGCTG	AGGCGCAGGAGCTTGGCAAG	1482
QY	301	CTGCAGATCGGACGGGCGCTGAGAACCTCTCA	CAGAGCCACAGAAACACCGGGTGAGA	360	QY	1381	GCACACTCAAGCGTCCCATCTCTGGGGCAGAGGAG	CGCGCTGCGGTGACAGCTGTGC	1440
Db	403	CTGCAGATCGGACGGGCGCTGAGAACCTCTCA	CAGAGCCACAGAAACACCGGGTGAGA	462	Db	1483	GCACACTCAAGCGTCCCATCTCTGGGGCAGAGGAG	CGCGCTGCGGTGACAGCTGTGC	1542
QY	361	GAGACGTTGGCCCTGGAGCTGAGTACCTCAACT	CGAAGGAGGAG	420	QY	1441	CGGCTCTGCGGAGGTGACCTGTTCGGGCTGTGAT	CTCCCGCAGAGCTGACGCTCA	1500
Db	463	GAGACGTTGGCCCTGGAGCTGAGTACCTCAACT	CGAAGGAGGAG	522	Db	1543	CGGCTCTGCGGAGGTGACCTGTTCGGGCTGTGAT	CTCCCGCAGAGCTGACGCTCA	1602
QY	421	CTGCAGAGCTCAGCGTGGCGTGGACCTTGGCGG	CATGGAGCGAAGCTGCATGTC	480	QY	1501	CTGGCCCAAGTATGCGGAGCTCGAACCTGAGGAT	GACTTCTGTGAGGCTGCCGAGGCCCG	1560
Db	523	CTGCAGAGCTCAGCGTGGCGTGGACCTTGGCGG	CATGGAGCGAAGCTGCATGTC	582	Db	1603	CTGGCCCAAGTATGCGGAGCTCGAACCTGAGGAT	GACTTCTGTGAGGCTGCCGAGGCCCG	1662
QY	481	CCCATGATCCCTTGGGCTGAGAGGAGACCAAG	AGCTGTGACTACCGCTGAAG	540	QY	1561	GACATCCAGCTTAAGACCCACCAAGAGCCAGAG	CGCGAGATGCCACGCTGTCCAGGGG	1620
Db	583	CCCATGATCCCTTGGGCTGAGAGGAGACCAAG	AGCTGTGACTACCGCTGAAG	642	Db	1663	GACATCCAGCTTAAGACCCACCAAGAGCCAGAG	CGCGAGATGCCACGCTGTCCAGGGG	1722
QY	541	GAGCTGATCTCAGTGCACTTTGGAGAGGAGCGG	CGCTCTCAGAGGAGAGAAATCAGGGAG	600	QY	1621	AAGGGGCTTGACATCTTCATTCGGCTGGGGCC	CTCTGTCTCAGCCAGAGAACCGG	1680
Db	643	GAGCTGATCTCAGTGCACTTTGGAGAGGAGCGG	CGCTCTCAGAGGAGAGAAATCAGGGAG	702	Db	1723	AAGGGGCTTGACATCTTCATTCGGCTGGGGCC	CTCTGTCTCAGCCAGAGAACCGG	1782
QY	601	CTGCAGGCTTGGGAGGAGCGGACCCCGAGCGA	ATGAGTGGGCTTGGAGCTG	660	QY	1681	TGGCGGCTGTGGGGCCGCTCACCTGACCCGAG	GAGAGGGCGCTTTGGCCCTCACGCTT	1740
Db	703	CTGCAGGCTTGGGAGGAGCGGACCCCGAGCGA	ATGAGTGGGCTTGGAGCTG	762	Db	1783	TGGCGGCTGTGGGGCCGCTCACCTGACCCGAG	GAGAGGGCGCTTTGGCCCTCACGCTT	1842
QY	661	CTCAGAGCTTATCAACACAGCTGTGCTTCTGAT	GCGGCTTCTCAACCCCTGCCAGG	720	QY	1741	CGGGGAGACTCGGCTGTCTCATCGCTGCGGT	CATTCCAGGGAGCCAGGCGCGGCGCT	1800
Db	763	CTCAGAGCTTATCAACACAGCTGTGCTTCTGAT	GCGGCTTCTCAACCCCTGCCAGG	822	Db	1843	CGGGGAGACTCGGCTGTCTCATCGCTGCGGT	CATTCCAGGGAGCCAGGCGCGGCGCT	1902
QY	721	AGCCTGGGCTTCTTCCACTGTGTAGCTGTCTT	ACTGGGGTCCCGGCGCCAGCAGCT	780	QY	1801	GGCCTGAAGAGGGCGACTACATTGTGTGAT	GTGAGGAGCCATCAGTGTGTGAGA	1860
Db	823	AGCCTGGGCTTCTTCCACTGTGTAGCTGTCTT	ACTGGGGTCCCGGCGCCAGCAGCT	882	Db	1903	GGCCTGAAGAGGGCGACTACATTGTGTGAT	GTGAGGAGCCATCAGTGTGTGAGA	1962
QY	781	GCCCTGGCTTCGAGAGGGGAGGCTTCTTCTT	CAACATGCGTGCCTCCACACGAGATT	840	QY	1861	CACGCGAGGTGTGTGACGAGCTCAAGGCTG	CGGGAGAGGCGGCGCCAGCTCAGGTG	1920
Db	883	GCCCTGGCTTCGAGAGGGGAGGCTTCTTCTT	CAACATGCGTGCCTCCACACGAGATT	942	Db	1963	CACGCGAGGTGTGTGACGAGCTCAAGGCTG	CGGGAGAGGCGGCGCCAGCTCAGGTG	2022
QY	841	GGGGCGGCGCAGGACCGCTCTGACCGAGGCTG	CGCGCGCTATGGAGGCTTCCAG	900	QY	1921	GTGTGCTGTGCCCCAGCTCTTAGACTGCC	AGCTTGGGGGAGCCCGCGCGCTCTGCTG	1980
Db	943	GGGGCGGCGCAGGACCGCTCTGACCGAGGCTG	CGCGCGCTATGGAGGCTTCCAG	1002	Db	2023	GTGTGCTGTGCCCCAGCTCTTAGACTGCC	AGCTTGGGGGAGCCCGCGCGCTCTGCTG	2082
QY	901	AGGGCGGCTGGGCTTCAGGCTCTCTGAGGAG	AACTTCTCCCATGCGCCGAGCCAGAC	960	QY	1981	GGCCCCAGGGGGCTTCTAAGAGGCGAGGG	AGCATGTTTCAAGACCCCGGCATCCACG	2040
Db	1003	AGGGCGGCTGGGCTTCAGGCTCTCTGAGGAG	AACTTCTCCCATGCGCCGAGCCAGAC	1062	Db	2083	GGCCCCAGGGGGCTTCTAAGAGGCGAGGG	AGCATGTTTCAAGACCCCGGCATCCACG	2142
QY	961	ATGAGCGCTGCGTCCCTCTGCGCACTGGAG	CAGCTCATGATGGCCCGAGGCCAGGAATGT	1020	QY	2041	TGGGCCAGTCCCGGGCCCTCTCTCAACTG	GAGCCGAAAGGCCCAAGGCGCAAGACTGGA	2100
Db	1063	ATGAGCGCTGCGTCCCTCTGCGCACTGGAG	CAGCTCATGATGGCCCGAGGCCAGGAATGT	1122	Db	2143	TGGGCCAGTCCCGGGCCCTCTCTCAACTG	GAGCCGAAAGGCCCAAGGCGCAAGACTGGA	2202
QY	1021	GTGTTTGAAGGCTCTACACCTGCTCCATGG	CCCCCCCCAAGTGTGCTGCGCCAGCTG	1080	QY	2101	GGTGTGCCCCAGGCTGTGCCCCCAGTGA	AGCAGCTCGGCCCTCATCTTCAAGCACCCA	2160
Db	1123	GTGTTTGAAGGCTCTACACCTGCTCCATGG	CCCCCCCCAAGTGTGCTGCGCCAGCTG	1182	Db	2203	GGTGTGCCCCAGGCTGTGCCCCCAGTGA	AGCAGCTCGGCCCTCATCTTGAAGCACCCA	2262
QY	1081	CGCCTGGCGCAGGAGCGCGCCAGGTGGCAG	CCGAGTACAGGCTAGTGACCCGACCATG	1140	QY	2161	GGGTGGCGG	2169	
Db	1183	CGCCTGGCGCAGGAGCGCGCCAGGTGGCAG	CCGAGTACAGGCTAGTGACCCGACCATG	1242	Db	2263	GGGTGGCGG	2271	
QY	1141	GCCAGCACCCGCTCCAGACTACGTTACGTT	GTCTCTGAGCTGCCCTGGTGCATGTCAAG	1200					
Db	1243	GCCAGCACCCGCTCCAGACTACGTTACGTT	GTCTCTGAGCTGCCCTGGTGCATGTCAAG	1302					

Db	1903	GGCCTGAAGAGGGGAGCTACATTGTGTGTCAGTGAATGGGAGCCATGCAGGTGGTGGAGA	1962
QY	1861	CACGCGGAGGTGTGACGGAGCTGAAGGCTGCGGGAGAGCGCGGCGCCAGCCTGCGAGGTG	1920
Db	1963	CACGCGGAGGTGTGACGGAGCTGAAGGCTGCGGGAGAGCGCGGCGCCAGCCTGCGAGGTG	2022
QY	1921	GTGTGCTGTGTCGCCAGCTCTAGACTGCCAGCTTGGGGGAGCGCGGCGCCGCTCTGCTG	1980
Db	2023	GTGTGCTGTGTCGCCAGCTCTAGACTGCCAGCTTGGGGGAGCGCGGCGCCGCTCTGCTG	2082
QY	1981	GGCCCCAGGGGGTCTTAAGAGGACGAGGAGCATGTTGCAAGACCCCGGCATCCACG	2040
Db	2083	GGCCCCAGGGGGTCTTAAGAGGACGAGGAGCATGTTGCAAGACCCCGGCATCCACG	2142
QY	2041	TGGGGCAGTCCCGGCCCTCTCAACTGAGGCGGAAAGGCCCGCAGGAGCAAGCTGGA	2100
Db	2143	TGGGGCAGTCCCGGCCCTCTCAACTGAGGCGGAAAGGCCCGCAGGAGCAAGCTGGA	2202
QY	2101	GGTGTCCCCAGCCCTGTGCCCCAGTGAAGCCAGCTCCGCCCTCATCTTTGAAGCACCCA	2160
Db	2203	GGTGTCCCCAGCCCTGTGCCCCAGTGAAGCCAGCTCCGCCCTCATCTTTGAAGCACCCA	2262
QY	2161	GGGTGGCCG	2169
Db	2263	GGGTGGCCG	2271
RESULT 3			
US-10-451-207-11			
; Sequence 11, Application US/10451207			
; Publication No. US20040038267A1			
; GENERAL INFORMATION:			
; APPLICANT: INCYTE GENOMICS, INC.			
; APPLICANT: BURFORD, Neil			
; APPLICANT: DING, Li			
; APPLICANT: YUE, Henry			
; APPLICANT: THORNTON, Michael B.			
; APPLICANT: CHAWLA, Narinder K.			
; APPLICANT: GANDHI, Aneena R.			
; APPLICANT: ARVIZU, Chandra S.			
; APPLICANT: BAUGHN, Mariah R.			
; APPLICANT: SWARNAKAR, Anita			
; APPLICANT: DUGGAN, Brendan M.			
; APPLICANT: LU, Dyung Aina M.			
; APPLICANT: THANGAVELU, Kavitha			
; APPLICANT: WARREN, Bridget A.			
; APPLICANT: TANG, Y. Tom			
; APPLICANT: KHAN, Farrah A.			
; APPLICANT: YAO, Monique G.			
; APPLICANT: EMBERLING, Brooke M.			
; TITLE OF INVENTION: INTRACELLULAR SIGNALING MOLECULES			
; FILE REFERENCE: PF-0868 USN			
; CURRENT APPLICATION NUMBER: US/10/451,207			
; PRIORITY FILING DATE: 2003-06-18			
; PRIOR APPLICATION NUMBER: PCT/US01/50315			
; PRIOR FILING DATE: 2001-12-19			
; PRIOR APPLICATION NUMBER: US 60/257,804			
; PRIOR FILING DATE: 2000-12-21			
; PRIOR APPLICATION NUMBER: US 60/260,102			
; PRIOR FILING DATE: 2001-01-05			
; NUMBER OF SEQ ID NOS: 18			
; SOFTWARE: PERL Program			
; SEQ ID NO 11			
; LENGTH: 2837			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
; FEATURE:			
; NAME/KEY: misc feature			
; OTHER INFORMATION: Incyte ID No. US20040038267A1 5560161CBI			
US-10-451-207-11			
Query Match 89.9%; Score 1951; DB 13; Length 2837;			
Best Local Similarity 100.0%; Pred. No. 0;			

```
QY 1299 CTTCTGTCAGCCGCCCACTCTCTTAAGCCCGAGGCCCTGTGCTCCCGAGAGAGCTGGA 1358
Db 1261 CTTCTGTCAGCCGCCCACTCTCTTAAGCCCGAGGCCCTGTGCTCCCGAGAGCTGGA 1320
QY 1359 GGAGCGCAGCAGCTTGCGCAAGGCACACCTGAAGCGTGCATCTCTGGGCGAGGAGGCG 1418
Db 1321 GGAGCGCAGCAGCTTGCGCAAGGCACACCTGAAGCGTGCATCTCTGGGCGAGGAGGCG 1380
QY 1419 GCTGGCGCTGCAAGCCCTGTGGCGCTCTGGCGAGGTGACCTCTTCGGGCTGTGAT 1478
Db 1381 GCTGGCGCTGCAAGCCCTGTGGCGCTCTGGCGAGGTGACCTCTTCGGGCTGTGAT 1440
QY 1479 CTTCCAGACGCTGCAGCGCTCACTGCCCAAGTATGGGAGCTGCACCGTGAAGATGACTT 1538
Db 1441 CTTCCAGACGCTGCAGCGCTCACTGCCCAAGTATGGGAGCTGCACCGTGAAGATGACTT 1500
QY 1539 CTGTGAGGCTGCCAGGCCGCCGGAATCCAGGCTTAAGACCCAGAGCCAGAGGCCAG 1598
Db 1501 CTGTGAGGCTGCCAGGCCGCCGGAATCCAGGCTTAAGACCCAGAGCCAGAGGCCAG 1560
QY 1599 GATGCCAGCCTGTCTCCAGGGGAAGGGGCTGACATCTTCCATCGGCTGGGCCCTGTGC 1658
Db 1561 GATGCCAGCCTGTCTCCAGGGGAAGGGGCTGACATCTTCCATCGGCTGGGCCCTGTGC 1620
QY 1659 TGTGTTCTCAGCAAGAACCGGTGGCGCTGTGGGCGCCGCTCCACCTGACCCGAGGAGA 1718
Db 1621 TGTGTTCTCAGCAAGAACCGGTGGCGCTGTGGGCGCCGCTCCACCTGACCCGAGGAGA 1680
QY 1719 GGGCGGCTTTGGCTTCAGCTTCGGGGAGACTCGCTGTCTCTCATCGCTGCCGTCAATCC 1778
Db 1681 GGGCGGCTTTGGCTTCAGCTTCGGGGAGACTCGCTGTCTCTCATCGCTGCCGTCAATCC 1740
QY 1779 AGGAGCCAGCCGCGCGGCTGGCTGAGAGGGGGAATCATTTGTGTCAAGTGAATGG 1838
Db 1741 AGGAGCCAGCCGCGCGGCTGGCTGAGAGGGGGAATCATTTGTGTCAAGTGAATGG 1800
QY 1839 GCAGCCATGCAAGTGTGGAGACACGCGAGGTGTGAGCGAGGTGTAAGGCTGCGGGAGA 1898
Db 1801 GCAGCCATGCAAGTGTGGAGACACGCGAGGTGTGAGCGAGGTGTAAGGCTGCGGGAGA 1860
QY 1899 GCGCGGCGCCAGCTGCAAGTGTGTGCTGCGCCAGCTCTAGACTGCCAGCTTCGG 1958
Db 1861 GCGCGGCGCCAGCTGCAAGTGTGTGCTGCGCCAGCTCTAGACTGCCAGCTTCGG 1920
QY 1959 GGAACGCGCGCGCTGCTGCTGGGCCCCCAGGGGGCTTCTAAGAGCCAGAGGAGCATGG 2018
Db 1921 GGAACGCGCGCGCTGCTGCTGGGCCCCCAGGGGGCTTCTAAGAGCCAGAGGAGCATGG 1980
QY 2019 TTGCAAGACCCCGGCATCCAGCTGGGCGCAGTCCCGCGCCCTCTCTCACTGGAGCCGAA 2078
Db 1981 TTGCAAGACCCCGGCATCCAGCTGGGCGCAGTCCCGCGCCCTCTCTCACTGGAGCCGAA 2040
QY 2079 GGGCCAGCAGGCGCAAGACTGGAGGCTGCCCCCAGCCCTGTGCCCCAGTGAAGCCAGCTCC 2138
Db 2041 GGGCCAGCAGGCGCAAGACTGGAGGCTGCCCCCAGCCCTGTGCCCCAGTGAAGCCAGCTCC 2100
QY 2139 GCCCTCATCTTTGAAGCACCCAGGGTGGCCG 2169
Db 2101 GCCCTCATCTTTGAAGCACCCAGGGTGGCCG 2131
```

RESULT 4

```
US-10-112-944-74
; Sequence 74, Application US/10112944
; Publication No. US20040048249A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Gezhi
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Xue, Aidong J.
```

```
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Wehrman, Tom
; APPLICANT: Ghosh, Malabika
; APPLICANT: Wang, Dunrui
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wang, Zhiwei
; TITLE OF INVENTION: No. US20040048249A1e1 Nucleic Acids and
; TITLE OF INVENTION: Secreted Polypeptides
; FILE REFERENCE: 805A
; CURRENT APPLICATION NUMBER: US/10/112,944
; PRIOR FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: US 09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 09/491,404
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: US 09/496,914
; PRIOR FILING DATE: 2000-02-03
; PRIOR APPLICATION NUMBER: US 09/515,126
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: US 09/519,705
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: US 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: US 09/552,929
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: US 09/577,408
; PRIOR FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 924
; SOFTWARE: pt_FL_genes Version 5.0
; SEQ ID NO 74
; LENGTH: 2738
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (103)..(2109)
US-10-112-944-74
```

Query Match 89.2%; Score 1935; DB 13; Length 2738;
Best Local Similarity 99.8%; Pred. No. 0;

Matches 1948; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

```
QY 219 GGGCTGTGACTCCCTGACGCGAGATCCAGTGGCGGCGAGCTGCAGAGCCGCGAGGCCAGAT 278
Db 162 GGGCTGTGACTCCCTGACGCGAGATCCAGTGGCGGCGAGCTGCAGAGCCGCGAGGCCAGAT 221
QY 279 TCACCAGCAGATTGACAGAGGAGCTGCAGATCGGACGGGCGCTGAGAACCTCTACAGAGC 338
Db 222 TCACCAGCAGATTGACAGAGGAGCTGCAGATCGGACGGGCGCTGAGAACCTCTACAGAGC 281
QY 339 CACCAGCAACACCGGGTGGAGAGACGGTCCGCTTGGAGCTGAGCTACGTCAACTCCAA 398
Db 282 CACCAGCAACACCGGGTGGAGAGACGGTCCGCTTGGAGCTGAGCTACGTCAACTCCA 341
QY 399 CTTGAGCTGCTGAAGAGGAGCTGGAGAGCTCAGCGGTGGCGTGAACCTTGGCCGCGCA 458
Db 342 CTTGAGCTGCTGAAGAGGAGCTGGAGAGCTCAGCGGTGGCGTGAACCTTGGCCGCGCA 401
QY 459 TGGGAGCGAAGCTGTCACTGTCCCATGATCCCCCTGGGCTTGAAGAGACCAAGAGAGCT 518
Db 402 TGGGAGCGAAGCTGTCACTGTCCCATGATCCCCCTGGGCTTGAAGAGACCAAGAGAGCT 461
QY 519 GGACTGCTTACACCGCTGAAGGAGCTGATCTCAGTGCACATTTTGGAGAGGACGGCGCTC 578
Db 462 GGACTGCTTACACCGCTGAAGGAGCTGATCTCAGTGCACATTTTGGAGAGGACGGCGCTC 521
QY 579 CTACGAGCGAGAATCAGGAGCTGGAGGCCCTTGGCGAGGCCCATGCGGACCCCGAGCCG 638
Db 522 CTACGAGCGAGAATCAGGAGCTGGAGGCCCTTGGCGAGGCCCATGCGGACCCCGAGCCG 581
QY 639 GAATGAGTCCGGCTGGAGCTGCTCAAGCCTATTACACAGCTGTGCTTCTTGATGTC 698
Db 582 GAATGAGTCCGGCTGGAGCTGCTCAAGCCTATTACACAGCTGTGCTTCTTGATGTC 641
```


QY 1615 CAGGGAGGGGCTGACATCTTCATCGCTGGGCGCCCTGTCTGTGTCTCAGCCAG 1674
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
373 CAGGGAGAGGGGCTGACATCTTCATCGCTGCCGCCCTGTCTGTGTCTCAGCCAG 432
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1675 AACCGGTGGCGGTGTGTGGGGCCGCTCCACCTGACCCGAGGAGGCGGCTTTGGCCCTC 1734
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
433 AACCGGTGGCGGTGTGTGGGGCCGCTCCACCTGACCCGAGGAGGCGGCTTTGGCCCTC 492
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1735 AGCTTTCGGGGAGACTCGGCTGTCTCATCGCTGCGCTCATTCAGGAGCCAGGCGCG 1794
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
493 AGCTTTCGGGGAGACTCGGCTGTCTCATCGCTGCGCTCATTCAGGAGCCAGGCGCG 552
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1795 GCGGCTGGCTCAAGGAGGCGGACTACATTTGTGTGAGTGAATGGGAGCCATGCAAGTGG 1854
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
553 GCGGCTGGCTCAAGGAGGCGGACTACATTTGTGTGAGTGAATGGGAGCCATGCAAGTGG 612
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1855 TGGAGACACGCGGAGGTGTGTGAGGAGCTGAAGGCTGCGGAGAGCGGCGCGCAGCCTG 1914
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
613 TGGAGACACGCGGAGGTGTGTGAGGAGCTGAAGGCTGCGGAGAGCGGCGCGCAGCCTG 672
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1915 CAGGTGGTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1974
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
673 CAGGTGGTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 732
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1975 CTGCTGGGCGCCAGAGGGGCTTCAAGGAGCCAGAGGAGCATGTTGCAAGACCCCGCA 2034
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
733 CTGCTGGGCGCCAGAGGGGCTTCAAGGAGCCAGAGGAGCATGTTGCAAGACCCCGCA 792
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2035 TCCAGTGGGCGCAGTCCCGGCGCCCTCTCAACTGGAGCCGAAAGCCCGCAGGCGCAAG 2094
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
793 TCCAGTGGGCGCAGTCCCGGCGCCCTCTCAACTGGAGCCGAAAGCCCGCAGGCGCAAG 852
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2095 ACTGAGGCTGCCCGCAGCCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2154
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
853 ACTGAGGCTGCCCGCAGCCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 912
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2155 CACCCAGGCTGGCCG 2169
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
913 CACCCAGGCTGGCCG 927

RESULT 6

US-10-023-896-15
; Sequence 15, Application US/10023896
; Publication No. US2003002776A1
; GENERAL INFORMATION:
; APPLICANT: Victor Roschke
; TITLE OF INVENTION: 29 Human Cancer Associated Proteins
; FILE REFERENCE: PA004PI
; CURRENT APPLICATION NUMBER: US/10/023,896
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: unassigned
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: PCT/US00/23794
; PRIOR FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/152,296
; PRIOR FILING DATE: 1999-09-03
; PRIOR APPLICATION NUMBER: 60/158,003
; PRIOR FILING DATE: 1999-10-06
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-023-896-15

Query Match 36.0%; Score 781.4; DB 15; Length 1013;
Best Local Similarity 99.7%; Pred. No. 3.6e-181;
Matches 793; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 1375 GGCAGGACACCTCAAGGCTGCCATCTTGGGGAGGAGGCGCTGGGCTGACGCC 1434
|||

Db 27 GGCAGGACACCTCAAGGCTGCCATCTTGGGGAGGAGGCGCTGGGCTGACGCC 86
QY 1435 CTGTGCGCGCTCTGTGCGAGGTGGACCTGTCTTGGGCTGTGATCTCCAGAGCGCTGCAG 1494
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
87 CTGTGCGCGCTCTGTGCGAGGTGGACCTGTCTTGGGCTGTGATCTCCAGAGCGCTGCAG 146
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1495 CGTCTCATCGGCAAGTATGCGAGGTGCGAGCTCGACCGTGTGAGATGACTTCTGTGAGGCTGCCAG 1554
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
147 CGTCTCACTGGCAGAGTATGCGAGGTGCGACCGTGTGAGATGACTTCTGTGAGGCTGCCAG 206
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1555 GCGCGGACATCCAGCTTAAGACCCACAGAGCCAGAGCCAGAGTGCACGCGCTGTCC 1614
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
207 GCGCGGACATCCAGCTTAAGACCCACAGAGCCAGAGCCAGAGTGCACGCGCTGTCC 266
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1615 CAGGGGAAAGGGGCTGACATCTTCCATCGGCTGGGGCCCTCTGTGTGTCTCAGCCAG 1674
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
267 CAGGGGAAAGGGGCTGACATCTTCCATCGGCTGGGGCCCTCTGTGTGTCTCAGCCAG 326
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1675 AACCGGTGGCGGTGTGTGGGGCCGCTCCACCTGACCCGAGGAGGCGGCTTTGGCCCTC 1734
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
327 AACCGGTGGCGGTGTGTGGGGCCGCTCCACCTGACCCGAGGAGGCGGCTTTGGCCCTC 386
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1735 AGCTTTCGGGGAGACTCGGCTGTCTCATCGCTGCGCTCATTCAGGAGCCAGGCGCG 1794
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
387 AGCTTTCGGGGAGACTCGGCTGTCTCATCGCTGCGCTCATTCAGGAGCCAGGCGCG 446
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1795 GCGGCTGGCTCAAGGAGGCGGCTTCAAGTGTGTGAGTGAATGGGAGCCATGCAAGTGG 1854
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
447 GCGGCTGGCTCAAGGAGGCGGCTTCAAGTGTGTGAGTGAATGGGAGCCATGCAAGTGG 506
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1855 TGGAGACACGCGGAGGTGTGTGAGGAGCTGAAGGCTGCGGAGAGCGGCGCGCAGCCTG 1914
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
507 TGGAGACACGCGGAGGTGTGTGAGGAGCTGAAGGCTGCGGAGAGCGGCGCGCAGCCTG 566
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1915 CAGGTGGTGTGCTGCTGCCAGCTTAGACTGCCAGCTTGGGGAGCCCGCGGCCCTC 1974
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
567 CAGGTGGTGTGCTGCTGCCAGCTTAGACTGCCAGCTTGGGGAGCCCGCGGCCCTC 626
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1975 CTGCTGGGCGCCAGGCGGCTTCAAGGAGCCAGAGGAGCATGTTGCAAGACCCCGCA 2034
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
627 CTGCTGGGCGCCAGGCGGCTTCAAGGAGCCAGAGGAGCATGTTGCAAGACCCCGCA 686
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2035 TCCAGTGGGCGCAGTCCCGGCGCCCTCTCAACTGGAGCCGAAAGCCCGCAGGCGCAAG 2094
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
687 TCCAGTGGGCGCAGTCCCGGCGCCCTCTCAACTGGAGCCGAAAGCCCGCAGGCGCAAG 746
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2095 ACTGAGGCTGCCCGCAGCCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2154
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
747 ACTGAGGCTG-CCCCAGCCCTGTGCCCCAGTGAAGCAGCTCCGCCCTCATCTTGAAG 805
QY 2155 CACCCAGGCTGGCCG 2169
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
806 CACCCAGGCTGGCCG 820

RESULT 7

US-09-925-301-174
; Sequence 174, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 174
; LENGTH: 1013

```
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-301-174

Query Match      36.0%; Score 779.8; DB 9; Length 1013;
Best Local Similarity 99.6%; Pred. No. 8.8e-181;
Matches 792; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1375 GGCAAGGCACACCTGAAGCGTGCATCTCGGGCAGGAGGCGCTGCGGCTGCACGCC 1434
Db 27 GGCAAGGCACACCTGAAGCGTGCATCTCGGGCAGGAGGCGCTGCGGCTGCACGCC 86

QY 1435 CTGTGCGCGCTCTCGCGAGGTGGACCTGCTTGGGCTGTGATCTCCAGACGCTGCAG 1494
Db 87 CTGTGCGCGCTCTCGCGAGGTGGACCTGCTTGGGCTGTGATCTCCAGACGCTGCAG 146

QY 1495 CGCTCACTGCCAAGTATGCGGAGCTCGACCGTGAAGATGACTTCTGTGAGGCTGCCGAG 1554
Db 147 CGCTCACTGCCAAGTATGCGGAGCTCGACCGTGAAGATGACTTCTGTGAGGCTGCCGAG 206

QY 1555 GCCCGGACATCCAGCCTTAAGACCCACAGAGCCAGAGCCAGGATGCCACGCTGTCC 1614
Db 207 GCCCGGACATCCAGCCTTAAGACCCACAGAGCCAGAGCCAGGATGCCACGCTGTCC 266

QY 1615 CAGGGGAAGGGCTGCATCTCCATCGGCTGGGCGCCCTGTCTGTGTTCTCAGCCAA 1674
Db 267 CAGGGGAAGGGCTGCATCTCCATCGGCTGGGCGCCCTGTCTGTGTTCTCAGCCAA 326

QY 1675 AACCGGTGGCGGCTGTGGGCGCCGTCACCTGACCCGAGAGGCGGCTTTGGCCTC 1734
Db 327 AACCGGTGGCGGCTGTGGGCGCCGTCACCTGACCCGAGAGGCGGCTTTGGCCTC 386

QY 1735 ACGTTTCGGGAGACTCGCTGTCTCATTCGCTCGCTCATTCAGGAGCCAGGCGCG 1794
Db 387 ACGTTTCGGGAGACTCGCTGTCTCATTCGCTCGCTCATTCAGGAGCCAGGCGCG 446

QY 1795 CGGCTCGCCTGAAGGAGGGCGACTACATTTGTGTCAGTGAATGGGAGCCATGCAAGTGG 1854
Db 447 CGGCTCGCCTGAAGGAGGGCGACTACATTTGTGTCAGTGAATGGGAGCCATGCAAGTGG 506

QY 1855 TGGAGACACGCGGAGGTGGTGAACGAGCTCGGAGAGGCGGCGCCAGCCTG 1914
Db 507 TGGAGACACGCGGAGGTGGTGAACGAGCTCGGAGAGGCGGCGCCAGCCTG 566

QY 1915 CAGTGTGTGCTGTGCTGCCAGCTCTAGACTGCCAGCTTGGGGGACCGCGCCGCTC 1974
Db 567 CAGTGTGTGCTGTGCTGCCAGCTCTAGACTGCCAGCTTGGGGGACCGCGCCGCTC 626

QY 1975 CTGCTGGGCGCCAGGGGCTTCTAAGAGCCAGAGGAGCATGTTGCAAGACCCCGCA 2034
Db 627 CTGCTGGGCGCCAGGGGCTTCTAAGAGCCAGAGGAGCATGTTGCAAGACCCCGCA 686

QY 2035 TCCACGTGGGCGAGTCCCGGCGCCCTCTCTCACTGGAGCCGAAAGGCCAGGCGCAAG 2094
Db 687 TCCACGTGGGCGAGTCCCGGCGCCCTCTCTCACTGGAGCCGAAAGGCCAGGCGCAAG 746

QY 2095 ACTGGAGGTCGCCCGGAGCCCTGTGCGGCTGAAGCCAGCTCGCGCTCATCTTTGAAG 2154
Db 747 ACTGGAGGTCGCCCGGAGCCCTGTGCGGCTGAAGCCAGCTCGCGCTCATCTTTGAAG 805

2155 CACCCAGGGTGGCCG 2169
806 CACCCAGGGTGGCCG 820
```

```
RESULT 8
US-10-023-896-43
; Sequence 43, Application US/10023896
; Publication No. US2003002776A1
; GENERAL INFORMATION:
; APPLICANT: Victor Roschke
; TITLE OF INVENTION: 29 Human Cancer Associated Proteins
; FILE REFERENCE: PA004P1
```

```
; CURRENT APPLICATION NUMBER: US/10/023,896
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: unassigned
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: PCT/US00/23794
; PRIOR FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/152,296
; PRIOR FILING DATE: 1999-09-03
; PRIOR APPLICATION NUMBER: 60/158,003
; PRIOR FILING DATE: 1999-10-06
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 43
; LENGTH: 1013
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-023-896-43

Query Match      36.0%; Score 779.8; DB 15; Length 1013;
Best Local Similarity 99.6%; Pred. No. 8.8e-181;
Matches 792; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1375 GGCAAGGCACACCTGAAGCGTGCATCTCGGGCAGGAGGCGCTGCGGCTGCACGCC 1434
Db 27 GGCAAGGCACACCTGAAGCGTGCATCTCGGGCAGGAGGCGCTGCGGCTGCACGCC 86

QY 1435 CTGTGCGCGCTCTCGCGAGGTGGACCTGCTTGGGCTGTGATCTCCAGACGCTGCAG 1494
Db 87 CTGTGCGCGCTCTCGCGAGGTGGACCTGCTTGGGCTGTGATCTCCAGACGCTGCAG 146

QY 1495 CGCTCACTGCCAAGTATGCGGAGCTCGACCGTGAAGATGACTTCTGTGAGGCTGCCGAG 1554
Db 147 CGCTCACTGCCAAGTATGCGGAGCTCGACCGTGAAGATGACTTCTGTGAGGCTGCCGAG 206

QY 1555 GCCCGGACATCCAGCCTTAAGACCCACAGAGCCAGAGCCAGGATGCCACGCTGTCC 1614
Db 207 GCCCGGACATCCAGCCTTAAGACCCACAGAGCCAGAGCCAGGATGCCACGCTGTCC 266

QY 1615 CAGGGGAAGGGCTGCATCTCCATCGGCTGGGCGCCCTGTCTGTGTTCTCAGCCAA 1674
Db 267 CAGGGGAAGGGCTGCATCTCCATCGGCTGGGCGCCCTGTCTGTGTTCTCAGCCAA 326

QY 1675 AACCGGTGGCGGCTGTGGGCGCCGTCACCTGACCCGAGAGGCGGCTTTGGCCTC 1734
Db 327 AACCGGTGGCGGCTGTGGGCGCCGTCACCTGACCCGAGAGGCGGCTTTGGCCTC 386

QY 1735 ACGTTTCGGGAGACTCGCTGTCTCATTCGCTCGCTCATTCAGGAGCCAGGCGCG 1794
Db 387 ACGTTTCGGGAGACTCGCTGTCTCATTCGCTCGCTCATTCAGGAGCCAGGCGCG 446

QY 1795 CGGCTCGCCTGAAGGAGGGCGACTACATTTGTGTCAGTGAATGGGAGCCATGCAAGTGG 1854
Db 447 CGGCTCGCCTGAAGGAGGGCGACTACATTTGTGTCAGTGAATGGGAGCCATGCAAGTGG 506

QY 1855 TGGAGACACGCGGAGGTGGTGAACGAGCTCGGAGAGGCGGCGCCAGCCTG 1914
Db 507 TGGAGACACGCGGAGGTGGTGAACGAGCTCGGAGAGGCGGCGCCAGCCTG 566

QY 1915 CAGTGTGTGCTGTGCTGCCAGCTCTAGACTGCCAGCTTGGGGGACCGCGCCGCTC 1974
Db 567 CAGTGTGTGCTGTGCTGCCAGCTCTAGACTGCCAGCTTGGGGGACCGCGCCGCTC 626

QY 1975 CTGCTGGGCGCCAGGGGCTTCTAAGAGCCAGAGGAGCATGTTGCAAGACCCCGCA 2034
Db 627 CTGCTGGGCGCCAGGGGCTTCTAAGAGCCAGAGGAGCATGTTGCAAGACCCCGCA 686

QY 2035 TCCACGTGGGCGAGTCCCGGCGCCCTCTCTCACTGGAGCCGAAAGGCCAGGCGCAAG 2094
Db 687 TCCACGTGGGCGAGTCCCGGCGCCCTCTCTCACTGGAGCCGAAAGGCCAGGCGCAAG 746

QY 2095 ACTGGAGGTCGCCCGGAGCCCTGTGCGGCTGAAGCCAGCTCGCGCTCATCTTTGAAG 2154
Db 747 ACTGGAGGTCGCCCGGAGCCCTGTGCGGCTGAAGCCAGCTCGCGCTCATCTTTGAAG 805
```

```
QY 2155 CACCCAGGTTGGCCG 2169
Db 806 CACCCAGGTTGGCCG 820

RESULT 9
US-10-112-944-558/c
; Sequence 558, Application US/10112944
; Publication No. US20040048249A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Gezhi
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Xue, Aidong J.
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Wehrman, Tom
; APPLICANT: Ghosh, Malabika
; APPLICANT: Wang, Dunrui
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wang, Zhiwei
; TITLE OF INVENTION: No. US20040048249A1el Nucleic Acids and
; TITLE OF INVENTION: Secreted Polypeptides
; FILE REFERENCE: 805A
; CURRENT APPLICATION NUMBER: US/10/112,944
; PRIOR FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: US 09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 09/491,404
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: US 09/496,914
; PRIOR FILING DATE: 2000-02-03
; PRIOR APPLICATION NUMBER: US 09/515,126
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: US 09/519,705
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: US 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: US 09/552,929
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: US 09/577,408
; PRIOR FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 924
; SOFTWARE: pt_FL_genes Version 5.0
; SEQ ID NO 558
; LENGTH: 1046
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(1046)
; OTHER INFORMATION: n = a,t,c or g
US-10-112-944-558

Query Match 24.2%; Score 524; DB 13; Length 1046;
Best Local Similarity 99.1%; Pred. No. 2.7e-118;
Matches 527; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1638 CCATCGCTGGGGCCCTCTGTGTCTTCAGCCAGAACCCGTCGGCGCTGGTGGGGCC 1697
Db 702 CCCCTGCCAGGGGCCCTCTGTGTCTTCAGCCAGAACCCGTCGGCGCTGGTGGGGCC 643

QY 1698 CGTCCACCTGACCCCGAGGAGAGGGCGCTTTGGCCCTCAGCTTTCGGGGAGACTCGCCTGT 1757
Db 642 CGTCCACCTGACCCCGAGGAGAGGGCGCTTTGGCCCTCAGCTTTCGGGGAGACTCGCCTGT 583

QY 1758 CTTATCGCTCGCTCATTCAGGGAGCCAGGCGCGGGCTGGCCCTGAAGAGGGCGA 1817
Db 582 CTTATCGCTCGCTCATTCAGGGAGCCAGGCGCGGGCTGGCCCTGAAGAGGGCGA 523

QY 1818 CTACATTGTTCAGTGAATGGGAGCCATGTCAGGTGGTGAGACACCGCGAGGTGGTGAC 1877
Db 522 CTACATTGTTCAGTGAATGGGAGCCATGTCAGGTGGTGAGACACCGCGAGGTGGTGAC 463
QY 1878 GGAGCTGAAGGCTGCGGGAGAGCGGGCCAGGCTTCAGAGTGGTGTGCTGCTGCCCCAG 1937
Db 462 GGAGCTGAAGGCTGCGGGAGAGCGGGCCAGGCTTCAGAGTGGTGTGCTGCTGCCCCAG 403
QY 1938 CTCTAGACTGCCAGCTTGGGGGACCGCGGGCCCGCTCTGTGGGCCCCAGGGGGCTTCT 1997
Db 402 CTCTAGACTGCCAGCTTGGGGGACCGCGGGCCCGCTCTGTGGGCCCCAGGGGGCTTCT 343
QY 1998 AAGGAGCCAGAGGAGCATGTTGCAAGACCCCGGCAATCCAGTGGGCGAGTCCCGGGCC 2057
Db 342 AAGGAGCCAGAGGAGCATGTTGCAAGACCCCGGCAATCCAGTGGGCGAGTCCCGGGCC 283
QY 2058 CCTCTCAACTGGAGCCGAAAGGCCAGCAGAGGCGCAAGACTGGAGGCTGCCCGGCCCTG 2117
Db 282 CCTCTCAACTGGAGCCGAAAGGCCAGCAGAGGCGCAAGACTGGAGGCTGCCCGGCCCTG 223
QY 2118 TGCCCCAGTGAAGCCAGCTCCGCCCTCATCTTGAAGCACCCAGGCGTGGCCG 2169
Db 222 TGCCCCAGTGAAGCCAGCTCCGCCCTCATCTTGAAGCACCCAGGCGTGGCCG 171

RESULT 10
US-10-239-607-52
; Sequence 52, Application US/10239607
; Publication No. US20030219761A1
; GENERAL INFORMATION:
; APPLICANT: Saatcioglu, Fahri
; TITLE OF INVENTION: No. US20030219761A1el Prostate-Specific or
; TITLE OF INVENTION: Testis-Specific Nucleic Acid Molecules, Polypeptides, and
; TITLE OF INVENTION: Diagnostic and Therapeutic Methods
; FILE REFERENCE: 50218/003002
; CURRENT APPLICATION NUMBER: US/10/239,607
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: PCT/US01/09410
; PRIOR FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 60/191,929
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 52
; LENGTH: 2100
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-239-607-52

Query Match 22.9%; Score 496.4; DB 16; Length 2100;
Best Local Similarity 56.8%; Pred. No. 1.5e-111;
Matches 977; Conservative 0; Mismatches 731; Indels 12; Gaps 3;

QY 219 GGGCTGTGACTCCCTGACGCGAGATCCAGTGGCGGCAGCTGCAGAGCCGCGAGGCCAGAT 278
Db 69 GGGCTGTGATCCCTTGCACAAACCGCGGAGTAATTCAGAAATCAAGAGCTGCTTT 128
QY 279 TCACCAGCAGATTCACAGAGAGCTGCAGATGGGCGCGCTGCAGAACCTCTACAGAGC 338
Db 129 GAATCAGCAGATTCCTGAAAGCGCTGCGGATGAGGACCGGAGCGGAAACCTTCTGAAAGT 188
QY 339 CACAGAGCAACACCGGCTGAGAGAGACCGTGCCTCTGAGAGTGCAGTCACTCA 398
Db 189 GGGCCAAACTCAAAAGGTGCGGGAGCAAGTGGGCTGGAGCTGAGCTTCTGCAACTCAGA 248
QY 399 CCTGCAGCTGCTGAAGGAGGAGCTGCAGAGAGCTCAGCGTGGCGTGGACCTTGGCGGCA 458
Db 249 CTTGCAGATGCTCAAGGAGAGCTGCAGAGGCTGAAACATCTCGTGGGCGTCTATCAGAA 308
QY 459 TGGGAGCGAAGCTGTCACTGTCTCCCATGATCCCCCTGGGCTTGAAGAGAGACCAAGAGCT 518
Db 309 CACAGAGGAGGACATTTACGATTTCCCTGATTCCTTTGGCCCTGAAGAGAAACGAAAGAGCT 368
QY 519 GGACTGGTCTACACCGCTGAGGAGAGCTGATCTCAGTGCACTTTTGGAGAGGAGCGGCGCTC 578
```


Db 369 CGACTTTGCGCTCTCAGGATTTATCTCGAACATTTACAGTGAAGATGCGTATTT 428
Qy 579 CTACGAGCAGAAATCAGGAGCTGGAGCCCTCGCGCAGGCCATCGGACCCCGAGCCG 638
Db 429 ATATGAAGATGAATTCAGATCTTATGATCTGAGACAAGCTTGTGCGAGCGCTAGCCG 488
Qy 639 GAATGAGTCCGGCTCGAGCTGCTCAGACGCTATTACAAACAGCTGTGCTTCTCGATGC 598
Db 489 GGATGAGCCCGGGTGGAACTGCTGATACATATCTCATCCAGCTGGCTTTGTCGAGAG 548
Qy 699 GCGCTTCTCTACCCCTCGCAGAGCCTCGGCTCTTCTTCCACTGGTACGACTCGCTTAC 758
Db 549 TCGAATCTTCCGCGCCACACGCGAGATGGACTCTCTTTCACCTGGTATGACTCTTCAC 608
Qy 759 TGGGGTCCCGCCCGCAGCAGCTGCTGGCCCTTCGAGAAAGGCGAGCTTCTTCTCAACAT 818
Db 609 TGGGGTTCGGTTCAGCCAGCAGAACTCTGCTGCGAGAAAGCCAGTGTCTGTTCACAC 668
Qy 819 CGGTGCTTCCACACGAGATTTGGGGCGCGCAGAGCCGCTCTGACCGAGGCTGCCG 878
Db 669 TGGGGCTCTTACACCCAGATTTGGGACCCCGTGGATCGGACGACGCTGGCTGGGA 728
Qy 879 CCGCGCTATGAGGCTTTCAGAGGGCGCTGGGGCCCTTCAGCCCTCTGAGGAGAACTT 938
Db 729 GAGTGCATAGATGCTTTCAGAGAGCGCAGGGGTTTAAATTTACCTGAAAGACACATT 788
Qy 939 CTCCTATGCGCGCAGCCAGACATGAGCGCTGGTCTCTCTGCGCATGGAGAGCTCAT 998
Db 789 TACCACATCTCAAGTTACGACTGAGCCCTGCCATGCTAGCGTGTCTGCTCAAAATGAT 848
Qy 999 GATGGCCAGGCCACAGGAATGTGTTTGGGGCTCTTCACACCTGCTCCATGGCCCC 1058
Db 849 GCTTGCACAGCCCAAGAAAGCTGTTTGAATAATCAGCTTCTCTG-----GATCCG 902
Qy 1059 CCAAGACTGCTGCGCCAGCTGCGCTGGCGCAGAGGCCGCCAGGTGGCAGCCGAGTA 1118
Db 903 GAATGAATTTCTATGCTGTGTGAAGTGGCTCAGAGGCTCTTAAGTGGGAGAGTCTA 962
Qy 1119 CAGCTAGTGCACCGGACCATGGCCAGCCAGCCCGTCCACACTACGTGCTGTCTCTG 1178
Db 963 CCAACAGCTACACGAGCCATGAGCCAGCGCCGCTGAAAGAAACATCCCTCTACTCCTG 1022
Qy 1179 GACTGCCCTGTGCTGATGCAAGGCGAGTACTTCGCTCTCCCTGCGCCACTACCACTAGC 1238
Db 1023 GGCAGCTTACCTGCTGAGGCGCCACCACTACGCGCCCTGCGCCACTTACTTCACTGC 1082
Qy 1239 CATGGCCCTCTGCGAGCGCTCCCGAGCGACCGAGGAGAGCTCCCGCAGCAGCAGGT 1298
Db 1083 CATCTCTCTATCGACACCAAGGTGAAGCCAGGCGAGATCTGGACACCAAGGAGAGTG 1142
Qy 1299 CTTCTCTGAGC-----CCCCACCTCTTAAGCCCGAGGCCCTGTGC--TGCGCGAGGA 1352
Db 1143 CCGTCCAGCTCTACGACCACTGCGAGAGGGGCTGACACCCCTTGGCCACACTGAAGAA 1202
Qy 1353 GCTGGAGGAGCGCAGGAGCTTGGCAGGCGACACTGAAGCGTGCATCTCTGGGCGAGGA 1412
Db 1203 TGATCAGAGCGCCGACAGCTGGGAAGTCCCACTTGGCAGAGCCATGGCTCATCAGGA 1262
Qy 1413 GGAGCGCTGCGGTGACGCGCTGTGCGCGTCTGCGGAGGTGAGCCTGCTTGGGCG 1472
Db 1263 GGAGTCCGTGCGGAGGCGAGCTCTGCAAGAAGCTGCGAGCATTTGAGGTGTACAGAA 1322
Qy 1473 TGTGATCTCCAGAGCTGAGCGCTCACTGCGCAAGTATCGGAGCTCGACCTGGAGGA 1532
Db 1323 GGTGCTGTGCGCGCAGGAACTCTCCCGCTCAGCTACGCCAGCAACAGGAGGAGGA 1382
Qy 1533 TGAATCTTGTGAGGCTGCGGAGGCCCGCGGACATCCAGCTTAAGACCCACAGAAAGCCAGA 1592
Db 1383 TGACTGCTGAACCTGATCGACGCGCCCGAGTGTGTTGCTTAAACTGAGCAAGAGTTGA 1442
Qy 1593 GGCAGGATGCCAGCTGTCCAGGGAAGGGGCTGACATCTTCCATCGGCTGGGCGC 1652

Db 1443 CATTATATTGCCCGAGTTTCCAAAGCTGACAGTCAAGGACTTCTTCCAGAAAGCTGGGCCC 1502
Qy 1553 CCTGCTGTGTTTCTCAGCCAAAGAACCGGTGCGGCTGGTGGGGCCCGTCCACCTGACCCG 1712
Db 1503 CTTATCTGTGTTTTCGGCTAACAGCGGTGAGCGCCTCTCGAAAGCATCCGCTTCACTGC 1562
Qy 1713 AGGAGAGGGGGCTTTGCGCTTCAGCTTCGCGGAGACTCGGCTCTCTCATCGCTGCCGT 1772
Db 1563 AGAAGAAGGGACATTTGGGTTTCACTTGAGAGGGAAGCCGCCCTTCAAGTTCATCTCT 1622
Qy 1773 CATTCCAGGGAGCCAGCCGCGCGCTGGCCCTGAAGAGGGCGCACTACATTTGTCTCAGT 1832
Db 1623 GATCTCTTACTGCTCTGCTCGGTGGCAGGAGCCCGGGAAGAGATTATATTGTCTCCAT 1682
Qy 1833 GAATGGCAGCCATGAGTGGTGGAGACACGCGAGGTGGTACGAGCTGAAGGCTGC 1892
Db 1683 TCAGCTTGTGATTTGAAGTGGCTGACGCTGAGTGAAGTTATGAAGCTGCTGAAGAGCTT 1742
Qy 1893 GGGAGAGCGGGCGCGCCAGCCCTGACAGTGGTGTCTGCTGCTG 1932
Db 1743 TGGCAGGACGAGATCGAGATGAAGTCTGTGAGCTCTCTG 1782

RESULT 11

US-10-239-607-51
; Sequence 51, Application US/10239607
; Publication No. US20030219761A1
; GENERAL INFORMATION:
; APPLICANT: Saatecioglu, Fahri
; TITLE OF INVENTION: No. US20030219761A1el Prostate-Specific or
; TITLE OF INVENTION: Testis-Specific Nucleic Acid Molecules, Polypeptides, and
; TITLE OF INVENTION: Diagnostic and Therapeutic Methods
; FILE REFERENCE: 50218/003002
; CURRENT APPLICATION NUMBER: US/10/239,607
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: PCT/US01/09410
; PRIOR FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 60/191,929
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 51
; LENGTH: 3526
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-239-607-51

Query Match 22.9%; Score 496.4; DB 16; Length 3526;
Best Local Similarity 56.8%; Pred. No. 1.5e-111;
Matches 977; Conservative 0; Mismatches 731; Indels 12; Gaps 3;

Qy 219 GGGCTGTGACTCCCTGAGCGAGATCCAGTGGCGCAGCTGCAGAGCCGAGGCCAGAT 278
Db 69 GGGCTGTAATCCCTTGCACAAACCGCGCGAGTAATTGCAGAAATCAAGAGCTGCTTT 128
Qy 279 TCACAGCAGATTGACAAAGGAGCTGCAGATCGGACCGCGGCTGAGAACCTCTTACAGAGC 338
Db 129 GAATCAGCAGATCTCTGAAGCCGTGCGATGAGGACCGGAGCGGAAACCTTCTGAAAGT 188
Qy 339 CACAGCAACAAACCGGCTGAGAGAGCGTCCCTGAGAGTGAAGTACGTTCACTCAATCAA 398
Db 189 GGCACAAACTCAAAAGGTGCGGAGCAAGTGGCGCTGAGCTGAGCTTCGTCAACTCAGA 248
Qy 399 CCTCAGCTGTGAAGGAGGAGCTGAGGAGCTCAGCGGTGGCTGAGCCCTGCGCCGCA 458
Db 249 CCTCAGATGCTCAAGGAGAGCTGAGAGGGCTGAAATCTCCGGTGGGCTTCTATCAGAA 308
Qy 459 TGGAGGGAAGCTGTCACTGTCCCATCATCCCCCTGGGCTGAAGAGACCAAGAGCT 518
Db 309 CACAGAGGAGGCAATTTACGATTCCTTCTTGGCTGAAGGAAACGAAAGAGCT 368
Qy 519 GGACTGTCTACACCGCTGAAGGAGCTGATCTCAGTGCATTTTGGAGAGAGCGGCGCTC 578

Db 369 CGACTTTGCGCTCGCTCAAGGATTTTATCTCGAACAATTACAGTGAAGATGGCTATT 428
QY 579 CTACGAGGCGAGAAATCAGGAGCTGAGAGCCCTGGCGGAGGCATCGGAGCCCGCAGCG 638
Db 429 ATATGAGATGAATTCAGATCTTATGATCTGAGCAAGCTTGTGCGAGCGCTAGCG 488
QY 639 GAATGAGTGGGCTGGAGCTGCTCACAGCTATTACAAACGAGTGTGTTCTCTGATGC 698
Db 489 GGATGAGGCGGGTGAATCTGATGACATACTTCACTCCAGCTGGGCTTTGTGAGAG 548
QY 699 GGGCTTCCTACCCCTGCGCAGAGCCTCGGCTCTCTTCCACTGCTGATGCTGCTTAC 758
Db 549 TCGATTTCTTCCGCGCCACAGCGAGATGGAGCTCTCTGTTCACTGATGATGCTCTCAC 608
QY 759 TGGGCTCCCGCCAGCAGCTGCTGCTTTCGAGAAAGGCGAGGTTCTCTTCAACAT 818
Db 609 TGGGCTTCGCTGAGCAGAGAACTGCTGCTGAGAGGCGAGTGTCTCTGTTCAACAC 668
QY 819 CGGTGCTTCCACAGCAGATTGGGCGCGCAGAGACCGCTCTGACCGAGGCTGCCG 878
Db 669 TGGGCGCTCTACACCCAGATTGGAGCCGCTGCGATCGGACAGCAGCGCTGGGCTGA 728
QY 879 CGCGCTATGGAGGCTTCCAGAGGCGCTGGGCTTCAGCTCTCGAGGAGAACTT 938
Db 729 GAGTGCCATAGATGCTTTCAGAGAGCGCGAGGGGTTTAAATTACTGAAAGACAT 788
QY 939 CTCCCATGCGCGAGCGCAGCATGAGCGCTGCTGCTGCGCACTGCGAGCAGCTCAT 998
Db 789 TACCCATATCTCAAGTTACGACATGAGCCCTGCATGCTCAGGCTGCTGTCAAATGAT 848
QY 999 GATGGCCCGAGGCGCAGGAATGTGTGTGAGGGCTCTCAACACCTGCTCCATGCGCCC 1058
Db 849 GCTTGCAAGCCCAAGAAAGCGTGTGTGAGAAATCAGGCTTCTCTG-----GGATCG 902
QY 1059 CBAAGACTGCTGGCCAGCTGCGCTGGCGAGGAGGCGCCAGCTGGCAGCCGAGTA 1118
Db 903 GAATGAATTTCTCATGCTGGTGAAGTGTGCTCAGGAGGCTGTAAAGTGGAGAGTCTA 962
QY 1119 CAGGCTAGTGCACCGACCATGCGCCAGCCAGCCCGCTCCAGCACTAGTGTGCTGCTCTG 1178
Db 963 CBAAGACTACAGCGAGCCATGAGCGAGCGCGCTGMAAGAGAAATCCCTTACTCTG 1022
QY 1179 GACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1238
Db 1023 GGCAGCTTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1082
QY 1239 CATGGGCTCTGCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1298
Db 1083 CATCTCTCTCATGCAACACAGGTGAAGCCAGGACCGATCTGGACACCGAGGAGTG 1142
QY 1299 CTTCTCTGCGAGC-----CCCCCAGCTCTCTTAAGCCCGGAGGCTGCTGCTGCTGCTG 1352
Db 1143 CTTCTCTGCGAGCTCTACAGCACATGCGAGAGGCTGACACCTTGGCCACACTGAAGAA 1202
QY 1353 GCTGAGAGGAGCGAGGAGCTTGGCAAGGACACCTGAGAGGCTGCTGCTGCTGCTGCTG 1412
Db 1203 TGATCAGCAGCGCGAGCTGCGGAGAGTCCCACTTGGCAGAGGCTGCTGCTGCTGCTG 1262
QY 1413 GAGGCGCTGCGGCTGCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1472
Db 1263 GAGTCTGCTGCGGAGCGAGCTCTCAAGAGCTGCGAGGCTGCGAGGCTGCTGCTGCTG 1322
QY 1473 TGTGATCTCCAGAGCTGCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1532
Db 1323 GGTGCTGCTGCGCAGAGAGCTCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1382
QY 1533 TGAATCTGCTGAGGCTGCGAGGCGCGGAGATCCAGGCTTAAGACCCACAGAGAGCGAGA 1592
Db 1383 TGACCTGCTGAACCTGATGAGCGCGCGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 1442
QY 1593 GGCAGGATGCGAGCTGCTGCGAGGAGGAGGCGCTGACATCTTCCATCGCTGCGGCG 1652
Db 1443 CATATATTTGCCCGAGTTCTCAAGCTGACAGTCAAGGAGTCTTCTCAGAGAGCTGCGG 1502

QY 1653 CTTGCTGCTGCTTCTCAGCCAAAGACCGGTGGCGGCTGGTGGGCGCGCTCCACTGACCG 1712
Db 1503 CTTATCTGTTGTTTTCGGCTTAAAGCGGTGAGCGCTCTCTCGAAGCATCCGCTTCACTG 1562
QY 1713 AGGAGGCGGCTTTGGCTTCAAGCTTTCGGGAGAGCTGCGCTGCTGCTGCTGCTGCT 1772
Db 1563 AGAAGAAGGGGACTTTGGGTTTCCCTTGGAGGGGAAACGCCCCGTTTCAAGTTTCACT 1622
QY 1773 CATTCCAGGAGCGGAGCGCGCGCTGCTGAGAGGAGGCGCACTACATTTGTTGCTGCT 1832
Db 1623 GATCTTACTGCTCTGCTTGGTGGAGAGCGCGGAGGAGATTAATTTGCTGCTCAT 1682
QY 1833 GAATGGGAGCGCATGAGGCTGCTGAGACACCGGAGAGTGTGACGAGAGCTGAGGCTG 1892
Db 1683 TCAGCTTGTGATTTGTAAGTGTGCTGAGCTGAGTGTGAGGTTATGAGCTGCTGAAGAG 1742
QY 1893 GAGAGGCGGCGCGCGCTGAGCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1932
Db 1743 TGGCAGGAGGAGATGAGATGAAAGTCTGAGAGCTTCTGCTGCTGCTGCTGCTGCTG 1782

RESULT 12

US-10-092-900A-351

; Sequence 351, Application US/10092900A

; Publication No. US20040043382A1

; GENERAL INFORMATION:

; APPLICANT: Padigaru, Muralidhara

; APPLICANT: Spytek, Kimberly A.

; APPLICANT: Shenoy, Suresh G.

; APPLICANT: Taupier Jr., Raymond J.

; APPLICANT: Pena, Carol E.A.

; APPLICANT: Li, Li

; APPLICANT: Zerhusen, Bryan D.

; APPLICANT: Gusev, Vladimir Y.

; APPLICANT: Ji, Weizhen

; APPLICANT: Gorman, Linda

; APPLICANT: Miller, Charles E.

; APPLICANT: Kekuda, Ramesh

; APPLICANT: Patturajan, Meera

; APPLICANT: Gangolli, Baha A.

; APPLICANT: Vernet, Corine A.M.

; APPLICANT: Guo, Xiaojia Sasha

; APPLICANT: Tchernev,, Velizar T.

; APPLICANT: Fernandes, Elma R.

; APPLICANT: Casman, Stacie J.

; APPLICANT: Malyankar, Uriel M.

; APPLICANT: Gerlach, Valerie

; APPLICANT: Liu, Yi

; APPLICANT: Anderson, David W.

; APPLICANT: Spaderna, Steven K.

; APPLICANT: Catterton, Elina

; APPLICANT: Leite, Mario W.

; APPLICANT: Zhong, Haibong

; APPLICANT: Alsobrook, John P.

; APPLICANT: Lepley, Denise M.

; APPLICANT: Rieger, Daniel K.

; APPLICANT: Burgess, Catherine E.

; TITLE OF INVENTION: No. US20040043382A1

; FILE REFERENCE: 21402-290C

; CURRENT APPLICATION NUMBER: US/10/092,900A

; PRIOR FILING DATE: 2002-03-07

; PRIOR APPLICATION NUMBER: USSN 60/274,322

; PRIOR FILING DATE: 2001-03-08

; PRIOR APPLICATION NUMBER: USSN 60/283,675

; PRIOR FILING DATE: 2001-04-13

; PRIOR APPLICATION NUMBER: USSN 60/338,092

; PRIOR FILING DATE: 2001-12-03

; PRIOR APPLICATION NUMBER: USSN 60/274,281

; PRIOR FILING DATE: 2001-03-08

; PRIOR APPLICATION NUMBER: USSN 60/274,191

; PRIOR FILING DATE: 2001-03-08

; PRIOR APPLICATION NUMBER: USSN 60/325,681

; APPLICANT: HELIX RESEARCH INSTITUTE			; FILE REFERENCE: H1-A0106		
; TITLE OF INVENTION: NO. US20040005560A1el full length cdna			; CURRENT APPLICATION NUMBER: US/10/108,260A		
; CURRENT FILING DATE: 2002-03-27			; NUMBER OF SEQ ID NOS: 5458		
; SOFTWARE: PatentIn Ver. 2.1			; SEQ ID NO 2397		
; LENGTH: 3500			; TYPE: DNA		
; ORGANISM: Homo sapiens			US-10-108-260A-2397		
Query Match 22.8%; Score 494.8; DB 16; Length 3500;			Best Local Similarity 56.7%; Pred. No. 3.7e-111;		
Matches 976; Conservative 0; Mismatches 732; Indels 12; Gaps 3;					
QY	219	GGGCTGTGACTCCCTGACCGCAGATCCAGTGGCGGCGAGCTGCAGAGCGCGCAGGCGCCAGAT	278		
DB	105	GGGCTGTAAATCCCTTGACAAACCGCGCGAGTAAATTCAGAAATCAAGAGCTGCTTT	164		
QY	279	TCACACAGAGATTGACAAAGGAGCTGCAGATGCGGCGCTGAGAACTCTACAGAGC	338		
DB	165	GAATCAGCAGATCTGAAAGCCATGCGATGAGACCGGAGCGGAACCTTCTGAAAGT	224		
QY	339	CACCAAGCAACCGGGTGAGAGACCGTGCCTCGGCTGAGCTAGCTACGTCACACTCAAA	398		
DB	225	GGCCACAAACTCAAAGGTGCGGAGCAAGTGGCGCTGGAGCTGAGCTTGGTCAACTCAGA	284		
QY	399	CTGCGAGTCTGAAAGGAGGAGCTGGAGAGCTCAGCGTGGCGTGGACCTTGGCGGCA	458		
DB	285	CCTGACAGATGCTCAAGGAAGAGCTGGAGGGCTGAACATCTCGGTGGCGCTTATCAGAA	344		
QY	459	TGGAGCGAAGCTGCACTGTCCCATGATCCCTCGGCGCTGAAGGAGCAACAGAGCT	518		
DB	345	CACAGAGGAGGATTAACGATTCCTGATTCCTTGGCTGAGGAACGAAAGAGCT	404		
QY	519	GGACTGGTCTACACCGCTGAAGAGCTGATCAGTGCACTTTGGAGAGGACGGCGCTC	578		
DB	405	CGACTTTGCGTGGTCTCAAGGATTTATCTGGAACTTACAGTGGAAGATGGCTATT	464		
QY	579	CTACAGGAGAAATCAGGAGCTGGAGCCCTGGGCGAGGCATGCGACCCCGAGCG	638		
DB	465	ATATGAAGATGAATTCAGATCTTATGGATCTGAGACAAAGCTTGTGCGAGCGCTAGCG	524		
QY	639	GAATCAGTGGCGCTGGAGCTGCTCACGCTATTACAAACAGCTGTGCTTCTTGATGC	698		
DB	525	GGATGAGCGCGGGTGGAACTGCTGATGACATCTTCATCCAGCTGGGCTTTGTGAGAG	584		
QY	699	GGCTTCTCTACCCCTGCGAGAGCTCGGGCTCTTTCCATGATGACACTGCTTAC	758		
DB	585	TCGATTTCTCCCGCCACACGCGCAGATGGACTCTCTGTTTCACTTGTATGACTCTCTAC	644		
QY	759	TGGGTCCCGGCCAGCAGCTGCTGCGCTTCGAGAGGCGAGGTTCTCTTCAACAT	818		
DB	645	CGGGTTCGGTTCAGCCAGCAACCTGCTGCTGGAGAGGCGAGTGTCTCTTCAACAC	704		
QY	819	CGGTGCGCTCCACACGCAATTGGGCGCGCGCAGAGACCGCTCTCTGACCGAGGCTGCCG	878		
DB	705	TGGGCGCTCTACACCCAGATTGGGACCCGCTGTGATCGGCAGACGCGGCTGGCTGA	764		
QY	879	CCGCGCTATGGAGGCTTCCAGAGGCGCTCGGGCTTTTTCAGCTTCTCGAGGAGAACTT	938		
DB	765	GAGTGCCATAGATGCTTTTCAGAGAGCGCGAGGGGTTTTAAATTAATCTGAAAGACAT	824		
QY	939	CTCCCATGCGGAGCCCGAGACATGAGCGCTGGTCCCTCTGGGCACTGGAGAGCTCAT	998		
DB	825	TACCCATCTCCAAAGTTACGACATGAGCCCTGCCATGCTCAGCGTCTCTGCAAAATGAT	884		
QY	999	GATGGCGCGAGCCAGGAATGTGTGTTTGGAGGCGCTCTCACACCTCGCTCCATGCGCCC	1058		
DB	885	GCTTGACAAACCCCAAGGAGGTGTTTGAGAAATCAGCGCTTCCCTG-----GGATCCG	938		

RESULT 14

US-09-895-040A-2
; Sequence 2, Application US/09895040A
; Patent No. US20020123474A1
; GENERAL INFORMATION:
; APPLICANT: Shannon, Mark
; TITLE OF INVENTION: HUMAN GTP-RHO BINDING PROTEIN 2
; FILE REFERENCE: AEOMICA-11
; CURRENT APPLICATION NUMBER: US/09/895,040A
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: PCT/US01/00666

QY	1059	CAAAGACTGCTGGCCCGAGCTGGCGCTGGCGCAGAGAGCGCGCCAGAGTGGCGAGCGAGTA	1118		
DB	939	GAATGAATTTCTCATGTGTGTGAAGTGGCTCAGGAGGCTGTAAAGTGGGAGAGGTCTTA	998		
QY	1119	CAGGCTAGTGACCGGACCATGGCCCGACCCCTCCACGACTACGTGCGCTGTCTCTCG	1178		
DB	999	CAACAGCTTACACGACCATGAGCCAGGCGCGGTGAAGAGAAACATCCCTACTCTCTG	1058		
QY	1179	GACTGCCCTGTGTCATGTCAAGCGCGAGTACTTTCGCTCCCTGGCCCACTACACGTAGC	1238		
DB	1059	GGCCAGCTTAGCTGCTGAAGGCCACCACTACGCGGCCCTGGCCCACTACTTCTCACTGC	1118		
QY	1239	CATGCGCTCTGCGACCGGTCCCGACGACCGAGGAGAGTCCCGACGACGAGAGAGT	1298		
DB	1119	CATCTCTCTCATCGACCAACCGAGTGAAGCCAGGACCGGATCTGGACCAACAGAGAAAGT	1178		
QY	1299	CTTCTCTGAGC-----CCCCCACTCTTAAGCCCGGAGGCCCTGTGC--TGCCGACGA	1352		
DB	1179	CTGTCCAGCTCTACGACCATGTCAGAGGGGTGACACCTTTGGCCACACTGAAGAA	1238		
QY	1353	GCTGAGGAGCGCAGCGAGCTTTGGCAAGGCACACCTGAAGCGTGCATCTCTGGGCGAGA	1412		
DB	1239	TGATCAGCAGCGCCGACAGCTGGGGAAGTCCCACTTTCGACAGAGCATGGCTCATCAGA	1298		
QY	1413	GGAGCGCTGCGGTGCGACGCTGCGCGCTGCGCGCTGCGCGAGGTGAGCTGCTTGGGCG	1472		
DB	1299	GGAGTCGGTGGCGGAGGCGAGCTCTGCAAGAAGCTGGCGAGCATTTGAGGTCTACAGAA	1358		
QY	1473	TGTGATCTCCACAGACGCTGCAGCGCTCACTGCGCAAGTATGCGGAGCTCGACCGTGAGA	1532		
DB	1359	GGTGTGTGTGCGCCGACAGGAACGCTCCCGCTACGTACGCGCCAGACACGAGAGAGGA	1418		
QY	1533	TGACTTCTGTGAGGCTGCGCGAGCGCCCGGACATTCAGCGCTTAAGCCCAACAGAGCCAGA	1592		
DB	1419	TGACCTGTGTAACCTGATCGACGCCCCCAGTGTGTGTGTAATACTGAGCAAGAGGTTGA	1478		
QY	1593	GGCCAGGATGCAACGCTGTCCAGGGGAAGGGGCTGACATCTTCATCGGCTGGGGCG	1652		
DB	1479	CATTATATTGCCCCAGTTCTCCAAGCTGACAGTACGCGACTTCTTCCAGAAAGCTGGGCGC	1538		
QY	1653	CCTGTCTGTGTCTCAGCCAGAACCGGTGGCGGTGGTGGGCGCGCTCCATCGCTGCCGT	1712		
DB	1539	CTTATCTGTGTTTTTTCGGCTAACAGCGTGGACGCTCTCTCGAAGCATCCGCTTCACTGC	1598		
QY	1713	AGGAGAGGCGGCTTTGGCTCACGCTTTCGGGAGACTGCGCTGCTCTCATCGCTGCCGT	1772		
DB	1599	AGAAGAAGGGACTTTGGGTTTCACTTGAGAGGAAACGCCCCGTTTCAAGTTCACTTCT	1658		
QY	1773	CATTCCAGGAGCGAGCGCGCGCTGGCTGAAGGAGGCGGAGCTACATTGTGTCTCAGT	1832		
DB	1659	GGATCTTACTGCTGCTGCTGCGTGGCAGGAGCCCGGAAAGAGATTATATTGTCTCCAT	1718		
QY	1833	GAATGGGAGCCATCAGGTGTGGAGACACGCGAGGTGGTGGCGAGCTGAAGGCTGAAGCTGC	1892		
DB	1719	TCAGCTTGTGATTTGTAAGTGTGCTGCGTGTGAGGTGTTTGAAGCTGCTGAAGAGCTT	1778		
QY	1893	GGGAGAGGCGGCGCGCAGCTGCGAGTGTGTGCTGCTGCTG 1932			
DB	1779	TGGCGAGGACGAGATCGAGATGAAGTCGTGAGCCTCTCTG 1818			

Db	1643	GGATCCTTACTGCTCTGCCCTCGGTGGCAGGAGCCCGGGAAGGAGATTATATTGTCTCCAT	1702
Qy	1833	GAATGGGCGCCATGCAGGTGGTGAGACACCGGGAGGTGGTCACGGAGCTGAAGGCTGC	1892
Db	1703	TCAGCTTGTGGATTGTAAGTGGCTGACGCTGAGTGAGGTTATGAAGCTGCTGAAGAGCTT	1762
Qy	1893	GGGAGAGCGGGCGCCAGCCTGCAGGTGGTGTCTGCTGCTG	1932
Db	1763	TGGCGAGGACGAGATCGAGATGAAGTCTGTGAGCCTCCTG	1802

Search completed: July 13, 2004, 12:02:44
Job time : 887.939 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 12, 2004, 20:25:00 ; Search time 3670.18 Seconds
(without alignments)
17647.909 Million cell updates/sec

Title: US-10-697-266-1_COPY_103_2271

Perfect score: 2169

Sequence: 1 agatcctgagagagggcc.....tgaagcaccaggggtggccg 2169

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estnu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_hcc:*

9: gb_est1:*

10: gb_est2:*

11: gb_hcc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_pmg:*

27: em_gss_vrl:*

28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
c 1	833.2	38.4	1070	13	BM802482
2	802.4	37.0	1001	13	BQ072325
3	796.4	36.7	1135	11	BC012135
4	787.6	36.3	899	13	BUI71128

c 5	761	35.1	1017	13	BM802482
c 6	738.6	34.1	1201	13	BM802482
c 7	722	33.3	1127	13	BM802482
8	716.6	33.0	1201	13	BM802482
9	715.8	33.0	1201	13	BM802482
c 10	708.6	32.7	1201	9	AL578313
c 11	624.4	28.8	1201	13	BM802482
c 12	613.2	28.3	1243	12	BM928275
c 13	605.6	27.9	679	10	BE251330
c 14	571	26.3	775	14	CA432432
c 15	556	25.6	1067	12	EG339711
c 16	544.4	25.1	3058	11	AK004849
c 17	541.2	25.0	3351	11	AK050214
c 18	533.6	24.6	830	12	BI518788
c 19	526	24.3	986	12	BI770906
c 20	515.6	23.8	1109	12	BM548005
c 21	513.2	23.7	793	12	BI103148
c 22	506	23.3	1201	9	AL555948
c 23	502.4	23.2	926	13	BQ880416
c 24	499.4	23.0	703	14	CB851722
c 25	495.4	22.8	707	14	CB850862
c 26	468.8	21.6	493	12	BM193145
c 27	468.6	21.6	709	10	BE408252
c 28	467.8	21.6	609	14	CB216237
c 29	464.4	21.4	494	12	BM153139
c 30	461	21.3	755	13	BQ179567
c 31	459.8	21.2	650	13	BU629871
c 32	458.6	21.1	810	10	BF607519
c 33	446	20.6	1408	12	BM475275
c 34	426	19.6	1021	13	BM444631
c 35	423.2	19.5	721	13	BQ442874
c 36	416	19.2	607	14	CD513476
c 37	411	18.9	955	13	BQ927119
c 38	404.8	18.7	612	9	AI917311
c 39	403	18.6	999	12	BM802482
c 40	401	18.5	875	13	BQ944101
c 41	386.8	17.8	793	13	BQ873217
c 42	385.2	17.8	671	12	BM743041
c 43	366	16.9	611	10	BE328691
c 44	356.8	16.4	539	12	BM839050
c 45	345.2	15.9	542	10	AW874410

ALIGNMENTS

RESULT 1
BX400896/c
LOCUS BX400896 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
DEFINITION cDNA clone CS0DK004YE05 3-PRIME, mRNA sequence.
ACCESSION BX400896.1 GI:30630225
VERSION BX400896.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1070)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 4751.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DK004AC03NP1&cluster=4751.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600

Faraday Avenue Genoscope sequence ID : CS0DK004AC03NP1.	
FEATURES	Location/Qualifiers
source	1..1070
	/organism="Homo sapiens"
	/mol_type="mRNA"
	/db_xref="taxon:9606"
	/clone="CS0DK004YE05"
	/cell_type="HELA"
	/cell_lines="HELA"
	/clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
	/note="1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN	
Query Match	38.4%; Score 833.2; DB 13; Length 1070;
Best Local Similarity	97.9%; Pred. No. 5.6e-134;
Matches 843; Conservative 9; Mismatches 8; Indels 1; Gaps 1;	
QY	1309 CCCCCACCTCTCTAAGCCCGAGCCCTGTGTCGCCAGAGCTGAGAGCGCAGG 1368
DB	974 SCCCKCACCTCTTAAACCCCGAGCCCTGT-YTCCCCAAGAGCTGAGAGCGCAGS 916
QY	1369 CAGCTTGGCAAGCACACTGAAGGTGCCATCTCGGGCAGAGGCGCTGGGCTG 1428
DB	915 CASCTTGGCAAGACACACTGAAGGTGCCATCTCGGGCAGAGGCGCTGGGCTG 856
QY	1429 CACGCCCTGTGCGGGTCTCGCGAGGTGACCTGCTTCGGGCTGTGATCTCCAGAGC 1488
DB	855 CACGCCCTGTGCGGGTCTCGCGAGGTGACCTGCTTCGGGCTGTGATCTCCAGAGC 796
QY	1489 CTGCAAGCTCACTGCGCAAGTATGCGAGCTCGACCGTGTGATGATCTGTGAGGCT 1548
DB	795 CTGCAAGCTCACTGCGCAAGTATGCGAGCTCGACCGTGTGAGGATGATCTGTGAGGCT 736
QY	1549 GCGAGGCCCGGACATCCAGCTTAAGCCACCAAGCCAGAGCCAGAGCCAGGATGCCAGC 1608
DB	735 GCGAGGCCCGGACATCCAGCTTAAGCCACCAAGCCAGAGCCAGAGGATGCCAGC 676
QY	1609 CTGTCCAGGGAAGGGGCTGACATCTTCCATCGGCTGGGGCCCTGTGTGTCTCA 1668
DB	675 CTGTCCAGGGAAGGGGCTGACATCTTCCATCGGCTGGGGCCCTGTGTGTCTCA 616
QY	1669 GCAAGAACCGGTGGCGCTGTGCGGCCCTGCTACCTGACCGAGGAGAGGGCGGCTTT 1728
DB	615 GCAAGAACCGGTGGCGCTGTGCGGCCCTGCTACCTGACCGAGGAGAGGGCGGNTTT 556
QY	1729 GGCCTCAGCTTCGGGAGACTCGCTGTCTCATCGCTGCGCTCATTCAGGAGCCAG 1788
DB	555 GGCCTCAGCTTCGGGAGACTCGCTGTCTCATCGCTGCGCTCATTCAGGAGCCAG 496
QY	1789 GCGCGCGCGCTGGCTGAAGAGGCGCACTACATTGTGTCACTGAATGGGAGCCCATGC 1848
DB	495 GCGCGCGCGCTGGCTGAAGAGGCGCACTACATTGTGTCACTGAATGGGAGCCATGC 436
QY	1849 AGTGTGTGAGACACCGCGAGTGTGACGAGCTGAAGGCTGCGGGAGAGCGGGCGCC 1908
DB	435 AGTGTGTGAGACACCGCGAGTGTGACGAGCTGAAGGCTGCGGGAGAGCGGGCGCC 376
QY	1909 AGCTCAGGTGGTGTGCTGCTGCCAGCTCTAGACTGCCAGCTTGGGGAGCCGCGG 1968
DB	375 AGCTCAGGTGGTGTGCTGCTGCCAGCTCTAGACTGCCAGCTTGGGGAGCCGCGG 316
QY	1969 CCGCTCTCTGGGCCCAAGGGGCTTCTAAGAGCCACAGGAGCATGGTTGCAAGACC 2028
DB	315 CCGCTCTCTGGGCCCAAGGGGCTTCTAAGAGCCACAGGAGCATGGTTGCAAGACC 256
QY	2029 CCGGCATCCAGTGGGCGCAGTCCCGGCCCTCTCACTGAGCGCGGAAGGCCAGCAG 2088
DB	255 CCGGCATCCAGTGGGCGCAGTCCCGGCCCTCTCACTGAGCGCGGAAGGCCAGCAG 196
QY	2089 GGCAAGACTGGAGGCTGCCCCAGCCCTGTGTGCCCGCAGTGAAGCCAGCTCCGCCCTCATCC 2148

Db	195	GGCAGACTGGAGGTGCCCGCCCTGTGCCCGAGTGGCCCGAGCTCCGCCCTCATCC	136
QY	2149	TTGAAGCACCCAGGCTGGCCG 2169	
Db	135	TTGAAGCACCCAGGCTGGCCG 115	
RESULT 2			
LOCUS	BQ072325	1001 bp	linear
DEFINITION	AGENCOURT 6838996 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5761460		EST 02-APR-2002
ACCESSION	BQ072325		5' mRNA Sequence.
VERSION	BQ072325.1	GI:19901371	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE	1 (bases 1 to 1001)		
AUTHORS	NIH-MGC http://mgc.hci.nih.gov/.		
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)		
JOURNAL	Unpublished (1999)		
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgabbs-remail.nih.gov Tissue Procurement: Life Technologies, Inc. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAM12810 row: b column: 21 High quality sequence stop: 623.		
FEATURES			
source		1..1001	
		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="IMAGE:5761460"	
		/lab_host="DH10B"	
		/clone_lib="NIH_MGC_122"	
		/note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source anonymous pool of 24 week female lung, 16 week female spleen, and 20-22 week male spleens. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 Kb, insert size range 1-3 Kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 026. Note: this is a NIH_MGC Library."	
ORIGIN			
Query Match		37.0%; Score 802.4; DB 13; Length 1001;	
Best Local Similarity		96.4%; Pred. No. 1.2e-128;	
Matches 864; Conservative 0; Mismatches 26; Indels 6; Gaps 4;			
QY	219	GGGCTGTGACTCCCTGACGATCCAGTTCAGTGGCGCCAGCTGCAGAGCCGAGGCCAGAT	278
DB	42	GGGCTGTGACTCCCTGACGATCCAGTTCAGTGGCGCCAGCTGCAGAGCCGAGGCCAGAT	101
QY	279	TCACAGCAGATTGCAAGGAGCTGCAGATCGGACGGCGCTGAGAACTCTACAGAGC	338
DB	102	TCACAGCAGATTGCAAGGAGCTGCAGATCGGACGGCGCTGAGAACTCTACAGAGC	161
QY	339	CACCAAGCAACCGGGTGCAGAGACCGTCCCTGGAGCTGAGTACCTCACTCCAA	398
DB	162	CACCAAGCAACCGGGTGCAGAGACCGTCCCTGGAGCTGAGTACCTCACTCCAA	221
QY	399	CTGCAAGCTGTGAAGGAGGAGCTGAGAGCTCAGCGTGGCTGGACCCCTGGCGGCA	458


```
Db 222 CCTGCAGCTGCTGAGGAGGAGCTGGAGAGCTGACGCGTGGCGCTGGACCTCGCGCGGCA 281
Qy 459 TGGAGGAGAACTGTCACTGTCCCATGATCCCTCGGCTGAGAGGAGACCAAGGAGCT 518
Db 282 TGGAGGAGAACTGTCACTGTCCCATGATCCCTCGGCTGAGAGGAGACCAAGGAGCT 341
Qy 519 GGACTGTCTACACCGCTGAGAGGAGCTGATCTCAGTGCACCTTGGAGAGGAGCGCGCTC 578
Db 342 GGACTGTCTACACCGCTGAGAGGAGCTGATCTCAGTGCACCTTGGAGAGGAGCGCGCTC 401
Qy 579 CTACGAGGCAAAATCAGGAGCTGGAGGCGCTCGGCGAGGCGCATGCGGACCCCGAGCG 638
Db 402 CTACGAGGCAAAATCAGGAGCTGGAGGCGCTCGGCGAGGCGCATGCGGACCCCGAGCG 461
Qy 639 GAATGAGTCGGGCTGGAGCTGTCTACAGGCTATTACAACAGCTGTGCTTCTTGGATGC 698
Db 462 GAATGAGTCGGGCTGGAGCTGTCTACAGGCTATTACAACAGCTGTGCTTCTTGGATGC 521
Qy 699 GCGCTTCTCACCCTGCCAGGAGCTCGGCTCTTCTTCCACTGGTACGACTCGCTTAC 758
Db 522 GCGCTTCTCACCCTGCCAGGAGCTCGGCTCTTCTTCCACTGGTACGACTCGCTTAC 581
Qy 759 TGGGCTCCCGCCAGCAGCTGCGCTTGGCTTGGAGGAGGCGCTTCTTCAACAT 818
Db 582 TGGGCTCCCGCCAGCAGCTGCGCTTGGCTTGGAGGAGGCGCTTCTTCAACAT 641
Qy 819 CGGTGCCCTCCACGAGATGGGGCGCGCAGGACCGCTTCTTCCAGGAGGCGCTTCTTCAACAT 878
Db 642 CGGTGCCCTCCACGAGATGGGGCGCGCAGGACCGCTTCTTCCAGGAGGCGCTTCTTCAACAT 701
Qy 879 CCGGCTATGAGGCTTCCAGAGGCGCGCTGGGCTTGGAGGAGGAGTCTTCAACAT 937
Db 702 CCGGCTATGAGGCTTCCAGAGGCGCGCTGGGCTTGGAGGAGGAGTCTTCAACAT 761
Qy 938 TCTCCCATGCGCG-AGCCAGACATGAGGCTCGCTTCTTCCAGGAGGAGGAGTCTTCAACAT 996
Db 762 TCTCCCATGCGCG-AGCCAGACATGAGGCTCGCTTCTTCCAGGAGGAGGAGTCTTCAACAT 821
Qy 997 ATGATGCGCCAGGAGGAGTGTGTGTG--AGGCGCTTCTTCCAGGAGGAGGAGTCTTCAACAT 1054
Db 822 ATGATGCGCCAGGAGGAGTGTGTGTG--AGGCGCTTCTTCCAGGAGGAGGAGTCTTCAACAT 881
Qy 1055 CCGCCAGAGCTCGCTG--CCAGCTGCGCTGCGCGAGGAGGAGGAGGAGTCTTCAACAT 1108
Db 882 CCGCCAGAGCTCGCTG--CCAGCTGCGCTGCGCGAGGAGGAGGAGGAGTCTTCAACAT 937

RESULT 3
BC012135
LOCUS
DEFINITION
Homo sapiens rhophilin, Rho GTPase binding protein 1, mRNA (cdna
clone IMAGE:4555079), with apparent retained intron.
BC012135
VERSION
Homo sapiens (human)
SOURCE
Homo sapiens
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1135)
Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
Stapleton M., Soares M.B., Bonaldo M.E., Casavant T.L.,
Schetz T.E., Brownstein M.J., Ustin T.B., Toshnycki S.,
Carninci P., Frange C., Raha S.S., Loquellano N.A., Peters G.J.,
Abramson R.D., Mullany S.J., Bosak S.A., McEwan P.J.,
McKernan K.J., Malek J.A., Gunaratne P.H., Richards S.,
Worley K.C., Hale S., Garcia A.M., Gay L.J., Huiyk S.W.,
Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Fahey J., Helton E., Kettman M., Madan A., Rodriguez S.,
```

```
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalls, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257
12477332
2 (bases 1 to 1135)
Strausberg, R.
Direct Submission
Submitted (02-AUG-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline
Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott,
Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/ILNL at: http://image.llnl.gov
Series: IRAL Plate: 29 Row: d Column: 17
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 19882236
This clone has the following problem: retained intron.
```

FEATURES

```
source
1. 1135
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4555079"
/tissue_type="Uterus, leiomyosarcoma"
/clone_lib="NIH MGC 46"
/lab_host="DH10B-R"
/note="Vector: pOTB7"

ORIGIN
```

```
Query Match 36.7% Score 796.4; DB 11; Length 1135;
Best Local Similarity 99.9%; Pred. No. 1.3e-127;
Matches 797; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy 219 GGGCTGTGACTCCCTGACGAGGAGCTGAGAGAGCGCTCGCCCTGGAGCTGAGCTCAACTCCAA 278
Db 162 GGGCTGTGACTCCCTGACGAGATTCAGTGGCGCCAGCTGCGAGCGCGCCAGAT 221
Qy 279 TCACGAGCAGATTGACAAAGGAGCTGAGAGAGCGCTCGCCCTGGAGCTGAGCTCAACTCCAA 338
Db 222 TCACGAGCAGATTGACAAAGGAGCTGAGAGAGCGCTCGCCCTGGAGCTGAGCTCAACTCCAA 281
Qy 339 CACGAGCAACAACCGGCTGAGAGAGAGCGCTCGCCCTGGAGCTGAGCTCAACTCCAA 398
Db 282 CACGAGCAACAACCGGCTGAGAGAGAGCGCTCGCCCTGGAGCTGAGCTCAACTCCAA 341
Qy 399 CTTGAGCTGTGAGAGGAGAGCTGAGAGAGCTCAGCGTGGCGTGGAGCCGCGGCA 458
Db 342 CTTGAGCTGTGAGAGGAGAGCTGAGAGAGCTCAGCGTGGCGTGGAGCCGCGGCA 401
```

QY	459	TGGGAGCGAAGCTGTCACTGTCTCCCATGATCCCTCGGCTTGAGGAGACCAAGAGCT	518
Db	402	TGGGAGCGAAGCTGTCACTGTCTCCCATGATCCCTCGGCTTGAGGAGACCAAGAGCT	461
QY	519	GGACTGTGCTACACCGCTGAAGGAGCTGATCTCAGTGCACCTTTGGAGAGGACGGGCTCTC	578
Db	462	GGACTGTGCTACACCGCTGAAGGAGCTGATCTCAGTGCACCTTTGGAGAGGACGGGCTCTC	521
QY	579	CTACGAGGCGAAGATCAGGAGCTGGAGCCCTGGCGCAGGCGCATGGGACCCCA3CGG	638
Db	522	CTACGAGGCGAAGATCAGGAGCTGGAGCCCTGGCGCAGGCGCATGGGACCCCA3CGG	581
QY	639	GAATGAGTGGGCTCGGAGCTGCTCACAGCCTATTACACACAGCTGTGCTTCTTGGATGC	698
Db	582	GAATGAGTGGGCTCGGAGCTGCTCACAGCCTATTACACACAGCTGTGCTTCTTGGATGC	641
QY	699	GGGCTTCTTCCACCCCTGCCAGGAGCTCGGGCTCTTCTTCCACTGGTACGACTCGCTTAC	758
Db	642	GGGCTTCTTCCACCCCTGCCAGGAGCTCGGGCTCTTCTTCCACTGGTACGACTCGCTTAC	701
QY	759	TGGGTFCCCGCCCGCAGCAGCTGGCCCTTGGCCTTCGAGAAAGGCGAGGTTCTCTTCAACAT	818
Db	702	TGGGTFCCCGCCCGCAGCAGCTGGCCCTTGGCCTTCGAGAAAGGCGAGGTTCTCTTCAACAT	761
QY	819	CGGTGCTTCCACACGACATTTGGGCGCGCCAGGACCGCTCTGACACGAGGCTCCCG	878
Db	762	CGGTGCTTCCACACGACATTTGGGCGCGCCAGGACCGCTCTGACACGAGGCTCCCG	821
QY	879	CGGCTGTATGAGGCGCTTCCAGAGGCGCTCGGGCTTTCAGCTCTCTGAGGAGAGACTT	938
Db	822	CGGCTGTATGAGGCGCTTCCAGAGGCGCTCGGGCTTTCAGCTCTCTGAGGAGAGACTT	881
QY	939	CTCCCATGCGCGAGCCCGCAGACATAGAGCTGGCTCCCTCTGGGCACTGGAGCAGCTCAT	998
Db	882	CTCCCATGCGCGAGCCCGCAGACATAGAGCTGGCTCCCTCTGGGCACTGGAGCAGCTCAT	941
QY	999	GATGGCCCGAGGCCAGGA	1016
Db	942	GATGGCCCGAGGCCAGGA	959
RESULT 4			
BUI71128			
LOCUS			
DEFINITION			
5', mRNA sequence.			
BUI71128			
BUI71128.1			
GI:22685112			
EST.			
Homo sapiens (human)			
Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo			
NIH-MGC			
http://mgi.nhl.gov/			
National Institutes of Health, Mammalian Gene Collection (MGC)			
Unpublished (1999)			
Contact: Robert Strausberg, Ph.D.			
Email: cgabbs-r@mail.nih.gov			
Tissue Procurement: ATCC			
CDNA Library Preparation: Life Technologies, Inc.			
DNA sequencing by: The I.M.A.G.E. Consortium (LLNL)			
Clone distribution: MGC clone distribution information can be			
found through the I.M.A.G.E. Consortium/LLNL at:			
http://image.llnl.gov			
Plate: L1AM13475			
High quality sequence stop: 563.			
Location/Qualifiers			
1. .899			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
FEATURES			
source			
RESULT 5			
BX444630/c			

/clone="IMAGE:6146655"			
/tissue type="retinoblastoma"			
/lab host="VDH10B (phage-resistant)"			
/clone_lib="NIH_MGC_87"			
/note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dr. Average insert size 1.75 kb. Library constructed by Life Technologies."			
ORIGIN			
Query Match			
Best Local Similarity			
Matches			
826; Conservative			
0; Mismatches			
9; Indels			
3; Gaps			
3; Length			
899;			
QY	1023	GTTCAGGCGCTCTCACACCTGCTCCATGTCGCCCCCAAGACTGCTTGCCAGCTGGG	1082
Db	1	GTTCAGGCGCTCTCACACCTGCTCCATGTCGCCCCCAAGACTGCTTGCCAGCTGGG	60
QY	1083	CCTGCGCAGGAGGCGCCAGGTGGCAGCCAGTACAGGCTAGTGCAACCGACCATGGC	1142
Db	61	CCTGCGCAGGAGGCGCCAGGTGGCAGCCAGTACAGGCTAGTGCAACCGACCATGGC	120
QY	1143	CAAGCCACCCGTCACAGCTAGTGGCTGTCTTGGACTGCCCTGGTGCANTGTCAAGC	1202
Db	121	CAAGCCACCCGTCACAGCTAGTGGCTGTCTTGGACTGCCCTGGTGCANTGTCAAGC	180
QY	1203	CGAGTACTTCCGCTCCCTGGCCCACTTACCACTAGCCATGGCCCTCTGCGACGGCTCC	1262
Db	181	CGAGTACTTCCGCTCCCTGGCCCACTTACCACTAGCCATGGCCCTCTGCGACGGCTCC	240
QY	1263	AGCGACCCGAGGAGAGCTCCCAACGACGAGCAGGTCTTCTGCGACCCCTCCTCTTC	1322
Db	241	AGCGACCCGAGGAGAGCTCCCAACGACGAGCAGGTCTTCTGCGACCCCTCCTCTTC	300
QY	1323	TAAGCCCGAGGCGCTGTGCTGCCGAGAGCTGGAGAGCGCAGCAGCTTGCGAAGGC	1382
Db	301	TAAGCCCGAGGCGCTGTGCTGCCGAGAGCTGGAGAGCGCAGCAGCTTGCGAAGGC	360
QY	1383	ACACCTGGAAGCTGCCATCTGCGGGCAGGAGAGGCGCTGCGGCTGCAGCCCTGTGCC	1442
Db	361	ACACCTGGAAGCTGCCATCTGCGGGCAGGAGAGGCGCTGCGGCTGCAGCCCTGTGCC	420
QY	1443	CGTCTTGGCGAGGTGAGCTGCTTCCGGCTGTGATCTCCAGACGCTGACGGCTCACT	1502
Db	421	CGTCTTGGCGAGGTGAGCTGCTTCCGGCTGTGATCTCCAGACGCTGACGGCTCACT	480
QY	1503	GGCCAAAGTATGGAGCTGACCGTGGAGTGAATCTTCTGAGGCTGCCGAGCCCGGA	1562
Db	481	GGCCAAAGTATGGAGCTGACCGTGGAGTGAATCTTCTGAGGCTGCCGAGCCCGGA	540
QY	1563	CATCCAGCCTAAGACCCACAGAACGACAGGCGCAGGATGCCACGGCTGTCCAGGGAA	1622
Db	541	CATCCAGCCTAAGACCCACAGAACGACAGGCGCAGGATGCCACGGCTGTCCAGGGAA	600
QY	1623	GGGG-CCTGACATCTTCCATTCGGCTGGGCGCTGCTGTGTTCTCAGCCAAAGACGGT	1681
Db	601	GGGG-CCTGACATCTTCCATTCGGCTGGGCGCTGCTGTGTTCTCAGCCAAAGACGGT	660
QY	1682	GGCGGCTGTGGGCGCGCTCCACTGACCGGAGAGAGGCGGCTTTGGCTCACGCTTC	1741
Db	661	GGCGGCTGTGGGCGCGCTCCACTGACCGGAGAGAGGCGGCTTTGGCTCACGCTTC	720
QY	1742	GGGGAGACTCGGCTGTCTCATCGCTGCCCTCATTTCCAGGAGCCAGGCGCGG-CGGCT	1800
Db	721	GGGGAGACTCGGCTGTCTCATCGCTGCCCTCATTTCCAGGAGCCAGGCGCGGCTTT	780
QY	1801	GGCTGAGAGGAGGCGCAGCTACATTT-GTGTGAGTGAATGGCAGCCATGACAGGTGGT	1857
Db	781	GGCTGAGAGGAGGCGCAGCTACATTTGGGCTCAGTGAATGGCAGCCATGACAGGGGGG	838

```

LOCUS      BX444630                      1017 bp      mRNA      linear      EST 15-MAY-2003
DEFINITION BX444630 Homo sapiens ADULT BRAIN Homo sapiens cDNA clone
ACCESSION  CS0DN004YH13 3-PRIME, mRNA sequence.
VERSION    BX444630
KEYWORDS   BX444630.1 GI:30778254
SOURCE     EST.
ORGANISM   Homo sapiens (human)

REFERENCE  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS    Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE      Full-length cDNA libraries and normalization
JOURNAL    Unpublished (2001)
COMMENT    Contact: Genoscope
           Genoscope - Centre National de Sequencage
           BP 191 91006 EVRY cedex - France
           Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
           Library was constructed by Life Technologies, a division of
           Invitrogen. This sequence belongs to sequence cluster 4751.f For
           more information about this cluster, see
           http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0DN004CD07NP1&cluster=4751.f. Contact :
           Feng Liang Email : fliang@lifetech.com URL :
           http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
           Faraday Avenue Genoscope sequence ID : CS0DN004CD07NP1.
           Location/Qualifiers
             1. .1017
               /organism="Homo sapiens"
               /mol_type="mRNA"
               /db_xref="taxon:9606"
               /clone="CS0DN004YH13"
               /tissue_type="ADULT BRAIN"
               /dev_stage="adult"
               /clone_lib="Homo sapiens ADULT BRAIN"
               /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
               was primed with a NotI-oligo(dT) primer. Five prime end
               enriched, double-strand cDNA was digested with Not I and
               cloned into the Not I and EcoRV sites of the pCMVSPORT 6
               vector. Library was not normalized."

FEATURES             source
ORIGIN
Query Match      35.1%; Score 761; DB 13; Length 1017;
Best Local Similarity 99.6%; Pred. No. 1.7e-121;
Matches 794; Conservative 0; Mismatches 0; Indels 3; Gaps 3;

QY 1375  GGCAAGGCACACCTGAAAGCGTGCATCTGGGGCAGGAGGCGCTGCGGCTGCAGGCC 1434
DB      |||
QY 1435  CTGTGCGCGCTCTGCGCGAGGTGGACCTGTTCCGGCTGTGATCTCCACAGCGCTGC 1492
DB      |||
QY 1493  AGCGTCACTGGCCCAAGTATGCGGAGCTCGACCTGAGGATGACTTCTGTAGGCTGCCG 1552
DB      |||
QY 1553  AGGCCCCGGGACATCCAGCTTAAGACCCACAGAGCCAGAGGCGCAGGATGCCAGCGCTGT 1612
DB      |||
QY 1613  CCNAGGGGAAGGGGCTTGAATCTTCCATCGGTGGGGCCCCCTGTCTGTGTCTTCAGCCA 1672
DB      |||
QY 1673  AGAACCGGTGGCGCTGGTGGGCGCCGCTCACTGACCCAGAGAGAGGGCGCTTTGGCC 1732
DB      |||
QY 1733  TCACGCTTCGGGGAGACTCGCCCTGTCTCTCATCGTCCGCTCATTCAGGAGGAGCGCCG 1792
DB      |||
QY 1793  CGGCGGCTGGCTGAAGGAGGCGGACTACATTGTCTCAGTGAATCGGAGCCATCGAGGT 1852
DB      |||
QY 1853  GGTGAGACACCGCGAGGTGGTGAACGAGCTGAAGCTTCGGGGAGAGGCGGCGCAGCC 1912
DB      |||
QY 1913  TGCAGGTGGTGTGCTGCTGCCAGCTTACAGTCCAGTCCAGCTTGGGGGAGCGCCGCCG 1972
DB      |||
QY 1973  TCCTGCTGGGCCCCCAGGGGGCTTCTAAGGAGCCAGAGGAGCATGTTGCAAGACCCCG 2032
DB      |||
QY 2032  TCCTGCTGGGCCCCCAGGGGGCTTCTAAGGAGCCAGAGGAGCATGTTGCAAGACCCCG 2092
DB      |||
QY 2093  CATCACGTGGCGCAGTCCCGGCCCTCTCAACTGGAGCCGAAAGCCCGCAGCAGGCA 2152
DB      |||
QY 2152  AGATGGAGGCTGCCCGAGCCCTGTGCCAGTGTGCCAGTGAAGCCA-CTCCGCCCTCATCCTTGA 160
QY 2153  AGCACCCAGGGGTGGCCG 2169
DB      |||
QY 2169  AGCACCCAGGGGTGGCCG 143
DB      |||

RESULT 6
LOCUS    BX447754/c
DEFINITION BX447754 Homo sapiens T CELLS (JURKAT CELL LINE) Homo sapiens cDNA
           clone CS0DH001YD09 3-PRIME, mRNA sequence.
ACCESSION BX447754
VERSION    BX447754.1 GI:31035860
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
AUTHORS    Full-length cDNA libraries and normalization
TITLE      Unpublished (2001)
JOURNAL    Contact: Genoscope
COMMENT    Genoscope - Centre National de Sequencage
           BP 191 91006 EVRY cedex - France
           Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
           was not normalized. Library was constructed by Life Technologies, a
           division of Invitrogen. This sequence belongs to sequence cluster
           4751.f For more information about this cluster, see
           http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0AH001CB05NP1&cluster=4751.f. Contact :
           Feng Liang Email : fliang@lifetech.com URL :
           http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
           Paraday Avenue Genoscope sequence ID : CS0AH001CB05NP1.
           Location/Qualifiers
             1. .1201
               /organism="Homo sapiens"
               /mol_type="mRNA"
               /db_xref="taxon:9606"
               /clone="CS0DH001YD09"
               /tissue_type="T CELLS (JURKAT CELL LINE)"
               /cell_line="JURKAT CELL LINE"
               /clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE)"
               /notes="Vector: pCMVSPORT 6; 1st strand cDNA was primed
               with a NotI-oligo(dT) primer. Five prime end enriched, into
               double-strand cDNA was digested with Not I and cloned, into
               the Not I and EcoRV sites of the pCMVSPORT 6 vector.
               Library was not normalized."

```

ORIGIN

Query Match		34.1%;	Score 738.6;	DB 13;	Length 1201;
Best Local Similarity		97.8%;	Pred. No. 1.4e-117;		
Matches 797;		Conservative 8;	Mismatches 8;	Indels 6;	Gaps 5;
QY	1357	GAGGAGCGCAGCAGCTTGGCAGGACACACCTGAAGCGTGCCATCTGGGGCAGAGGAG	1416		
DB	976	GAAGAGCGCAGCAGCTTGGCAAG--MACACCTRAGCGTGCCATCTYGGGGCAGAGGAG	919		
QY	1417	GGCTGTGGGCTGCACGCC--TGTCGCGCTCTTGGCGAGGTGACCTCTTGGGCTGT	1475		
DB	918	GGCTGTGGGCTGCACGCCCTTGTGCGCGCTCTTGGCGAGGTGACCTCTTGGGCTGT	859		
QY	1476	GATCTCCAGACGCTCAGCGCTCATCTGCGCAAGTATGCGGAGCTCGACCGGTGAGCATCA	1535		
DB	858	GATCTCCACGA--GCTGCAGCGCTCACTGGCCAAATATGCGGAGCTCGACCGGTGAGCATCA	800		
QY	1536	CTTCTGTGAGGCTGCCGAGGCCCGGACATCCAGCCTAAGACCCACCAAGAGCCACAGGC	1595		
DB	799	CTTCTGTGAGGCTGCCGAGGCCCGGACATCCAGCCTAAGACCCACCAAGAGCCACAGGC	740		
QY	1596	CAGGATGCCACGCTGTCTCCAGGGGAAGGGGCTGACATCTTCCATCGGCTGGGGCCCT	1655		
DB	739	CAGGATGCCACGCTGTCTCCAGGGGAAGGGGCTGACATCTTCCATCGGCTGGGGCCCT	680		
QY	1656	CTCTGTGTTCTCAGCCAGAACCGGTGGCGGCTGTGGGCGCGTCCACTGACCCGAG--	1714		
DB	679	GTCGTGTGTTCTCAGCCAGAACCGGTGGCGGCTGTGGGCGCGTCCACTGACCCGAG--	620		
QY	1715	GAGAGGGCGGCTTTGGCCCTCAGCCTTCGGGGAGACTCGCGCTGTCCCTCATCGCTGCC	1774		
DB	619	GAGAGGGCGGCTTTGGCCCTTACGCTTCGGGGAGACTCGCGCTGTCCCTCATCGCTGCC	560		
QY	1775	TTCCAGGGAGCCAGGCCGCGGCGGCTGGGCTTGAAGGAGGGCGACTACATGTGTCTGTGA	1834		
DB	559	TTCCAGGGAGCCAGGCCGCGGCGGCTGGGCTTGAAGGAGGGCGACTACATGTGTCTGTGA	500		
QY	1835	ATGGGAGCCATGACAGTGTGTGAGACACGCGAGGTGTGTGCGGAGCTGAAGGC	1894		
DB	499	ATGGGAGCCATGACAGTGTGTGAGACACGCGAGGTGTGTGCGGAGCTGAAGGC	440		
QY	1895	GAGAGGGCGGCGGCGGCTGCAGGTGTGTGCTGTGCGGAGCTCTAGACTGCCAGCT	1954		
DB	439	GAGAGGGCGGCGGCGGCTGCAGGTGTGTGCTGTGCGGAGCTCTAGACTGCCAGCT	380		
QY	1955	TGGGGAGCCGCGGCGGCTGTCTGTGGGCCCGGAGGGGCTTCTAAGGAGCCAGAGGAGC	2014		
DB	379	TGGGGAGCCGCGGCGGCTGTCTGTGGGCCCGGAGGGGCTTCTAAGGAGCCAGAGGAGC	320		
QY	2015	ATGGTTGCAAGACCCCGCATCCAGTGGGCCAGTCCCGGCGGCTCTCAACTGTAGCC	2074		
DB	319	ATGGTTGCAAGACCCCGCATCCAGTGGGCCAGTCCCGGCGGCTCTCAACTGTAGCC	260		
QY	2075	GAAAGGCCAGCGGCGGAGCTGAGGTGTCCCGGAGGCTGTGCGGCGGAGTGAAGCCAG	2134		
DB	259	GAAAGGCCAGCGGCGGAGCTGAGGTGTGCGGAGGCTGTGCGGCGGAGTGAAGCCAG	200		
QY	2135	CTCGGCTCATCTTTGAAGCACCCAGGTGGCG	2169		
DB	199	-TCCGCGCTCATCTTTGAAGCACCCAGGTGGCG	166		

RESULT 7
BX341691
LOCUS
DEFINITION BX341691 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
cDNA clone CS0DK004YE05 5-PRIME, mRNA sequence.
ACCESSION BX341691
VERSION BX341691
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Euthera; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 1127)
Li W.B., Gruber C., Jessee J. and Polayes D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 4751.f For
more information about this cluster, see
http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0DK004AC03QPl
&cluster=4751.f. Contact: Feng Liang Email: fliang@lifetech.com
URL: http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID: CS0DK004AC03QPl.
Location/Qualifiers
1..1127
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DK004YE05"
/cell_type="HELA CELLS COT 25-NORMALIZED"
/cell_line="HELA"
/clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match		33.3%;	Score 722;	DB 13;	Length 1127;
Best Local Similarity		95.5%;	Pred. No. 1e-114;		
Matches 806;		Conservative 10;	Mismatches 20;	Indels 8;	Gaps 7;
QY	219	GGGCTGTGACTCCCTGACCGAGATCCAGTGGCGGAGCTGCAGAGCGGCGGAGGCGGCGGAGAT	278		
DB	267	GGGCTGTGACTCCCTGACCGAGATCCAGTGGCGGAGCTGCAGAGCGGCGGAGGCGGCGGAGAT	325		
QY	279	TCACACGAGATTGACAGGAGCTGCAGATCGGAGCGGCGGCTGAGAACCTCTACAGAGC	338		
DB	326	TCACACGAGATTGACAGGAGCTGCAGATCGGAGCGGCGGCTGAGAACCTCTACAGAGC	385		
QY	339	CACGAGCAACAAACCGGCTGAGAGAGCGTGCCTTGGAGCTGAGTACGTCAACTCAA	398		
DB	386	CACGAGCAACAAACCGGCTGAGAGAGCGTGCCTTGGAGCTGAGTACGTCAACTCAA	445		
QY	399	CTTGAGCTGTGAAGGAGCTGAGAGAGCTCAGGGTGGCGTGGACCTGCGCGGCA	458		
DB	446	CTTGAGCTGTGAAGGAGCTGAGAGAGCTCAGGGTGGCGTGGACCTGCGCGGCA	505		
QY	459	TGGGAGCGAGCTGTCACTGTCCCATGATCCCTTGGGCTGAAGGAGACCAAGAGCT	518		
DB	506	TGGGAGCGAA--CTGTCACTGTCCCATGATCCCTTGGGCTGAAGGAGACCAAGAGCT	564		
QY	519	GGACTGTCTACACCGCTGAAGGAGCTGATCTCAGTGCACTTTGGAGAGACGCGGCTC	578		
DB	565	GGACTGTCTACACCGCTGAAGGAGCTGATCTCAGTGCACTTTGGAGAGACGCGGCTC	624		
QY	579	CTACGAGGAGAAATCAGGAGCTGAGGCGGCTGCGGAGGCGCATGCGACCCCGAGCGG	638		
DB	625	CTACGAGGAGAAATCAGGAGCTGAGGCGGCTGCGGAGGCGCATGCGACCCCGAGCGG	684		
QY	639	GAATGAGTCGGGCTGGAGCTGCTCAAGCTTATTACACAGCTGTGCTTCTGGATGC	698		
DB	685	GAATGAGTCGGGCTGGAGCTGCTCAAGCTTATTACACAGCTGTGCTTCTGGATGC	744		
QY	699	GCGCTTCTCACCCTGCGGAGGCTTCTTCCACTGTACACTGCTGCTTAC	758		
DB	745	GCGCTTCTCACCCTGCGGAGGCTTCTTCCACTGTGCTGCTTAC	804		
QY	759	TGGGTCCTCCGCGGCGGAGCGCTGCTGCGCTTCGAGAGGCGGAGCTTCTTCAACAT	818		

Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 4751.f For more information about this cluster, see

<http://www.genoscope.cns.fr/>

<http://cluster.cgi?seq=CS0AH001CB05QP1&cluster=4751.f>. Contact : Feng Liang Email : fliang@lifetech.com URL :

<http://fulllength.invitrogen.com/> Invitrogen Corporation 1600

Faraday Avenue Genoscope sequence ID : CS0AH001CB05QP1.

FEATURES

source

1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DH001YD09"
/tissue_type="T CELLS (JURKAT CELL LINE)"
/cell_line="JURKAT CELL LINE"
/clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE)"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

ORIGIN

Query Match 33.0%; Score 715.8; DB 13; Length 1201;
Best Local Similarity 91.9%; Pred. No. 1.2e-113;
Matches 824; Conservative 6; Mismatches 13; Indels 54; (aps 5;
QY 227 ACTCCCTGACGAGATCCAGTGGCGCAGCTGCAGAGCCGAGCGGCCAGATTCA/CAGC 286
Db 211 ACCCGGGCTGCAGATCCAGTGGCGCAGCTGCAGAGCCGCA-GGCCAGATTCA/CAGC 269
QY 287 AGATTGCAAGAGCTGCAGATCGGACGGGGCTGAGAACCTCTACAGAGCCAC/AGCA 346
Db 270 AGATTGCAAGAGCTGCAGATCGGACGGGGCTGAGAACCTCTACAGAGCCAC/AGCA 329
QY 347 ACAACCGGTGAGAGAGCGTCCCTGGAGCTGAGTCACTCAACTCAACTC/CAGC 406
Db 330 ACAACCGGTGAGAGAGCGTCCCTGGAGCTGAGTCACTCAACTCAACTC/CAGC 389
QY 407 TGCTGAAGGAGAGCTGGAGAGCTCAGCGGTGGCGTGGACCTTGCCTGCGGATGG/AGCG 466
Db 390 TGCTGAAGGAGAGCTGGAGAGCTCAGCGGTGGCGTGGACCTTGCCTGCGGATGG/AGCG 449
QY 467 AGCTGTCACTGCCCATGATCCCTGGCGCTGAGGAGACACAGGAGCTGGA/TTGT 526
Db 450 AA-CTGTCACTGCCCATGATCCCTGGCGCTGAGGAGACACAGGAGCTGGA/TTGT 508
QY 527 CTACACCGCTGAAGGAGCTGATCTCAGTGCATTTGGAGAGAGCGGCGCTTCTTA/GAGG 586
Db 509 CTACACCGCTGAAGGAGCTGATCTCAGTGCATTTGGAGAGAGCGGCGCTTCTTA/GAGG 568
QY 587 CAGAAATCAGGAGCTGGAGGCGCTCGGCGAGCCATTCGCGACCCCGAGCGGAA/GAGT 646
Db 569 CAGAAATCAGGAGCTGGAGGCGCTCGGCGAGCCATTCGCGACCCCGAGCGGAA/GAGT 628
QY 647 CGGGCTGGAGCTGCACAGCTATTACACAGCTGTCTTCTCTGGAGCGCG/TTCC 706
Db 629 CGGGCTGGAGCTGCACAGCTATTACACAGCTGTCTTCTCTGGAGCGCG/TTCC 688
QY 707 TCACCCCTGCAGAGAGCTCGGGCTCTTCTTCCACTGGTACGACTCGCTTACTGG/TTCC 766
Db 689 TCACCCCTGCAGAGAGCTCGGGCTCTTCTTCCACTGGTACGACTCGCTTACTGG/TTCC 748
QY 767 CGGCCACAGAGCTGCGCTTGGAGAGGCGAGCTTCTTCAACATCGG/TTCC 826
Db 749 CGGCCACAGAGCTGCGCTTGGAGAGGCGAGCTTCTTCAACATCGG/TTCC 808
QY 827 TCACACGCGAGATT-GGGGCGCGCAGAGCGCTCTTGTGACCGAGGCTGCGCGC/TTCC 885
Db 809 TCACACGCGAGATTGGGGGCGCGCAGAGCGCTCTTGTGACCGAGGCTGCGCGC/TTCC 868
QY 886 ATGAGAGCGCTTCCAGAGAGGCGCTGGGGCTTTCAGCGCTCTGTGAGGAGAACTTCT/CCAT 945

Db 859 ATGGAGGCGCTTCCAGAGGCGCGCT----- 892
QY 946 GGGCCGAGCCACAGATGAGCGCTGGTCCCTCTGGCACTGGAGCAGCTCATGATGGCC 1005
Db 893 -----GACATGAGCGCTGGTCCCTCTGGCACTGGAGCAGCTCATGATGGCC 940
QY 1006 CAGGCCCGAGGAATGTGTGTTGAGGGCTCTTCCACACCTTCCATGCCATGCCCGCCCAAGAC 1065
Db 941 CAGGCCCGAGGAATGTGTGTTGAGGGCTCTTCCACACCTTCCATGCCATGCCCGCCCAAGAC 1000
QY 1066 TCCCTGGCCAGCTGCGCTGGCGCAGGAGCGGCCAGGTGGCGCCGAGTACAGG 1122
Db 1001 TGCTTGGCCAGCTGCGC---TGGSGRGAAGCGCCAGTGGCGAGTACAGS 1054
RESULT 10 1201 bp mRNA linear EST 01-JUN-2003
AL578313 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
DEFINITION CDNA clone CS0DK010YM04 3-PRIME, mRNA sequence.
ACCESSION AL578313 GI:31316527
VERSION EST.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1201)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On Feb 16, 2001 this sequence version replaced gi:12942264.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seq@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 4751.f For more information about this cluster, see <http://www.genoscope.cns.fr/>
<http://cluster.cgi?seq=CS0DK010BG02NP1&cluster=4751.f>. Contact : Feng Liang Email : fliang@lifetech.com URL : <http://fulllength.invitrogen.com/> Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CS0DK010BG02NP1.
FEATURES
source
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DK010YM04"
/cell_type="HELA CELLS COT 25-NORMALIZED"
/cell_line="HELA"
/clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."
ORIGIN
Query Match 32.7%; Score 708.6; DB 9; Length 1201;
Best Local Similarity 95.7%; Pred. No. 2.2e-112;
Matches 761; Conservative 8; Mismatches 22; Indels 4; Gaps 4;
QY 1376 GCAAGGCACACCTGAAAGCGTGCATCTCTGGGGCAGGAGGCGCTGCGCTGACGCCCC 1435
Db 946 KCTGTGCCCCASTSGYTTGCGAGGAGCGCCACAGGAGGCGCTGC-GCTGACGCCCC 888
QY 1436 TGTGCGCGTCTTCCGCGAGGTGGAAGCTTCTTGGGCTGTGATCTTCCAGCAGCTGAGC 1495
Db 887 TGTGCGCGTCTTCC-GCGGAGGTGGAAGCTTCTTGGGCTGTGATCTTCCAGCAGCTGAGC 829
QY 1496 GCTCACTGGCCAAAGTATGCGAGCTCGACCGTGAGGATGACTTCTGTGAGGCTGCCGAGG 1555

<hr/>					
BM928275.1 GI-19378654					
EST.					
VERSION	Homo sapiens (human)				
KEYWORDS	Homo sapiens				
SOURCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
ORGANISM	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
<hr/>					
REFERENCE	1 (bases 1 to 1243)				
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/.				
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)				
JOURNAL	Unpublished (1999)				
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgaabs-f@mail.nih.gov Tissue Procurement: Life Technologies, Inc. CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov				
<hr/>					
FEATURES	Plate: LHAM2832 row: k column: 06 High quality sequence start: 115 High quality sequence stop: 630. Location/Qualifiers 1..1243 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:5770109" /lab_host="DH10B" /clone_lib="NIH_MGC_121" /note="Organ: brain; Vector: pCMV-SPORE; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source anonymous pool of 3 fetal brains, female age 20 weeks, female age 24 weeks, and male age 26 weeks. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.7 kb, insert size range 0.7-3.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 017. Note: this is a NIH_MGC Library."				
<hr/>					
ORIGIN					
Query Match	28.3%; Score 613.2; DB 12; Length 1243;				
Best Local Similarity	90.8%; Pred. No. 7.2e-96;				
Matches	655; Conservative 0; Mismatches 63; Indels 3; Gaps 1;				
<hr/>					
QY	758	CTGGGTCCCGGCCACGAGCGTGCCCTGGCCTTCGAGAGGGCAGCGTTCTTCAACA	817		
Dd	27	CTGNCCCCCNCCCACACGNGCCCTTGGCCTTCAACAANGNCATCGTTCTTCAACA	86		
<hr/>					
QY	818	TCGGTGCCCTCCACACGAGATTGGGGCGGCCACGAGCACCGCTCCTGCACCGAGGTCGCC	877		
Dd	87	TCNATGCCCTCCACAGCAATTGNGGGCGCCANGACCACCTCTGCACCCAGGTCGCC	146		
<hr/>					
QY	878	GCCCGGTATGAGGCGCTTCAGAGGGCGCTGGGGCGCTTCAGCGCTCTCAGGGAGAACT	937		
Dd	147	GCCGCGTATGAGGCGCTTCAGAGGGCGCTGGGGCGCTTCANCTCTCAGGGACAAC	206		
<hr/>					
QY	938	TTCTCCCATGGCCGAGCCACACATGAGCGCTGCGTCCCTTGAGGCCCTCTACCACTGC	997		
Dd	207	TTCTCCCATGGCCGAGCCAAACATGAACGCTGCGTCCCTCTGCGCACTGGAGCAN	266		
<hr/>					
QY	998	TGATGGCCCGAGCCAGGAATGTGTGTTTGAAGGCCCTCTACCACTGCCTCCATGGCC	1057		
Dd	267	TGATGGCCCGAGCCAGGAATGTGTGTTTGAAGGCCCTCTACCACTGCCTCCATGGCC	326		
<hr/>					
QY	1058	CCCAGACTGCCTGGCCAGCTGCCTGGCGCAGAGGGCGGCCAGGTGCGACCGCAGT	1117		
Dd	327	CCCAGACTGCCTGGCCCANCTGCGCCTTGGCGCAGAGGGCGGCCAGGTGCGACCGCAGT	386		
<hr/>					
QY	1118	ACAGGTAGTCACGGGACCATGGCCACGCCACCGCTCCAGCTAGCTGCTGTCTCT	1177		
Dd	387	ACAGGTAGTCACGGGACCATGGCCACGCCACCGCTCCAGCTAGCTGCTGTCTCT	446		


```
Db      237  CTGTGCCCCAGTGAAGCCAGCTCGGCCCTCATCTTTGAAGCACCACCCAGGCTGGCG 183
|||||
RESULT 15
BG339711
LOCUS   602437292F1 NIH_MGC_46 Homo sapiens cDNA clone IMAGE:4555079 5',
DEFINITION mRNA sequence.
ACCESSION BG339711
VERSION   BG339711.1 GI:13146149
KEYWORDS EST.
SOURCE    Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1067)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ARCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCMI255 row: h column: 24
High quality sequence stop: 704.
Location/Qualifiers
1..1067
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4555079"
/tissue_type="leiomyosarcoma cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_46"
(note="Organ: uterus; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCACGAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
library.")
ORIGIN
Query Match 25.6%; Score 556; DB 12; Length 1067;
Best Local Similarity 86.7%; Pred. No. 5.4e-86;
Matches 671; Conservative 0; Mismatches 95; Indels 8; Gaps 5;
QY 219 GGGCTGTGACTCCCTGACGAGATCCAGTGGCGGCAGCTGCGAGCGCGAGGCCGCCAGAT 278
Db 162 GGGCTGTGACTCCCTGACGAGATCCAGTGGCGGCAGCTGCGAGCGCGAGGCCGCCAGAT 221
QY 279 TCACGACGAGATTGACAAGGAGCTGCAGATGGGACGGCGCTGAGAACTCTACGAGC 338
Db 222 TCACGACGAGATTGACAAGGAGCTGCAGATGGGACGGCGCTGAGAACTCTACGAGC 281
QY 339 CACGACCAACACCGGGTGAGAGAGCGTCCCTTGGAGCTGAGCTAGCTCAACTCCAA 398
Db 282 CACGACCAACACCGGGTGAGAGAGCGTCCCTTGGAGCTGAGCTAGCTCAACTCCAA 341
QY 399 CTGCGAGCTGCTGAAGAGGAGCTGAGGAGCTGACGCGTGGCGTGGACCCCTGGCGGCA 458
Db 342 CTGCGAGCTGCTGAAGAGGAGCTGAGGAGCTGACGCGTGGCGTGGACCCCTGGCGGCA 401
QY 459 TGGGACGAGCTGTCACTGTCCCATGATCCCTGGCCCTGAAGGAGCAAGGAGCT 518
|||||
```

```
Db      402 TGGGAGCGAAGCTGTCACTGTCCCCATGATCCCCCTGGGCTGAAGGAGACCAAGGAGCT 461
QY 519 GGACTGGTCTACACCGCTGAAGGAGCTGATCTCAGTGCACCTTTGGAGAGAGCGGCCCTC 578
Db 462 GGACTGGTCTACACCGCTGAAGGAGCTGATCTCAGTGCACCTTTGGAGAGAGCGGCCCTC 521
QY 579 CTACGAGGCGAGAAATCAGGGAGCTGGAGCCCTGGGGCAGGCCATGCGGACCCCCAGCCG 638
Db 522 CTACGAGGCGAGAAATCAGGGAGCTGGAGCCCTGGGGCAGGCCATGCGGACCCCCAGCCA 581
QY 639 GAATGAGTCGGCCCTGGAGCTCTCACAGCCTATTACAACACAGCTGTGCTTCTTGATGC 698
Db 582 GAATGAGTCGGCCCTGGAGCTCTCACAGCCTATTACAACACAGCTGTGCTTCTTGATGC 641
QY 699 GCGCTTCTCTCACCCCTGCCAGGAGCTCGGGCTCTTCTTCCACTGTACGACTCGCTTAC 758
Db 642 GCGCTTCTCAGCCCTGCGAGAGCCTCGGGCTCTTCTTCCACTGTGTTACGACTCGCTTAC 701
QY 759 TGGGGTCCCGGCCCGACGAGCGTGGCCCTTCGAGAAAGGCGAGCGTTCTCTTCAACAT 818
Db 702 TGGGGTCCCGG--CGAGGAGCGTGCCTTG--CTTCGAGAAGGCGCGGTCTCTTCAACAT 759
QY 819 CCGTGCCTTCCACACGACAGATTGGGGCGCGCCAGGACCGCTCTCTGACCCGAGGCTCCCG 878
Db 760 CCGTGCCT--TCCAAGGCGAGATTGGGGCGCGCCAGGACCGTCTCTTGAAGTGTGCCCG 817
QY 879 CCGCGCTA--TGGAGGCGCTTCCAGAGGCGCGCT---GGGCGCTTCAAGCTCTTGAGGAGA 934
Db 818 CCGCGCTATTGAGGGTTTCCAGAGGGGCGTTGGGGGCTTTGGGCTTCTTGAGGGGGA 877
QY 935 ACTTCTCCCATGCGCGAGCCCGACAGACATGAGCGCTGCGTCCCTCTGCGCACTGG 988
Db 878 GCTTCTCCCGTTGGGCGGCGCAGGAATTAGGGGCGTGGGTCTTATGGCCTGG 931
|||||
```

Search completed: July 13, 2004, 08:21:02
Job time : 3673.68 secs

Db	463	GAGACGTCGCCCTGGAGCTGAGCTACGTCAACTCCAACTCGACGCTGCTGAGGAGGAG	522
Qy	421	CTGGAGAGCTCAGCGTGGCGTGGACCTCTGGCCGCATGSGAGCGAAGCTGTCACTGTC	480
Db	523	CTGGAGAGCTCAGCGTGGCGTGGACCTCTGGCCGCATGSGAGCGAAGCTGTCACTGTC	582
Qy	481	CCCATGATCCCTTGGGCTGAGAGGACCAAGGAGCTGSACTGCTTACACCGCTGAAG	540
Db	583	CCCATGATCCCTTGGGCTGAGAGGACCAAGGAGCTGSACTGCTTACACCGCTGAAG	642
Qy	541	GAGCTGATCTCAGTGCACTTTGGAGAGGACGGCGCCTCTTACGAGGCGAGAAATCAGGGAG	600
Db	643	GAGCTGATCTCAGTGCACTTTGGAGAGGACGGCGCCTCTTACGAGGCGAGAAATCAGGGAG	702
Qy	601	CTGAGGCCCCTGGGCGAGGCGATCGGACACCCACAGCCGGAATGAGTCGGGCCCTGAGCTG	660
Db	703	CTGAGGCCCCTGGGCGAGGCGATCGGACACCCACAGCCGGAATGAGTCGGGCCCTGAGCTG	762
Qy	661	CTCACAGCTATTACAAACAGCTGTGCTTCTGTGATGCGCTTCTTCAACCTCGCCAGG	720
Db	763	CTCACAGCTATTACAAACAGCTGTGCTTCTGTGATGCGCTTCTTCAACCTCGCCAGG	822
Qy	721	AGCCTCGGGCTCTTTCTTCCATGTTACGACTCGCTTACTCGGGTCCCGGCCACGACGCT	780
Db	823	AGCCTCGGGCTCTTTCTTCCATGTTACGACTCGCTTACTCGGGTCCCGGCCACGACGCT	882
Qy	781	GCCCTGGCTTCGAGAGGGCAGCGTTCTTTCAACATCGTGGCCCTCCACACGCAATT	840
Db	883	GCCCTGGCTTCGAGAGGGCAGCGTTCTTTCAACATCGTGGCCCTCCACACGCAATT	942
Qy	841	GGGCGCGCCAGGACCGCTCTCTGACACGAGGGTCCCGCGCGCTATGAGGCGCTTCCAG	900
Db	943	GGGCGCGCCAGGACCGCTCTCTGACACGAGGGTCCCGCGCGCTATGAGGCGCTTCCAG	1002
Qy	901	AGGCGCGCTGGGGCTTCAGCCTCTCTGAGGAGAACTTCTCCATGCGCGCGACCCAGC	960
Db	1003	AGGCGCGCTGGGGCTTCAGCCTCTCTGAGGAGAACTTCTCCATGCGCGCGACCCAGC	1062
Qy	961	ATGAGCGCTGCGTCCCTCTGCGACTGAGCAGCTCATGATGGCCACGAGGCCCAGGAATGT	1020
Db	1063	ATGAGCGCTGCGTCCCTCTGCGACTGAGCAGCTCATGATGGCCACGAGGCCCAGGAATGT	1122
Qy	1021	GTGTTTGAAGGCCCTCTCACACCTGCTCCATGCCCCCACAAGACTGCTGGGCCAGCTG	1080
Db	1123	GTGTTTGAAGGCCCTCTCACACCTGCTCCATGCCCCCACAAGACTGCTGGGCCAGCTG	1182
Qy	1081	CGCTGGCGCAGAGGCGCGCCAGGTGGCAGCCAGGTACAGGTAGTGCAACGACCATG	1140
Db	1183	CGCTGGCGCAGAGGCGCGCCAGGTGGCAGCCAGGTAGTGCAACGACCATG	1242
Qy	1141	GCCAGCACCCGCTCCACGACTAGTGCTCTCTCTGGACTGCGCTTGGTGTCAATCAAG	1200
Db	1243	GCCAGCACCCGCTCCACGACTAGTGCTCTCTCTGGACTGCGCTTGGTGTCAATCAAG	1302
Qy	1201	GCGAGTACTTCGCGTCCCTGGCCCACTACAGTAGGCATGGGCCCTTGGAGCGGCTCC	1260
Db	1303	GCGAGTACTTCGCGTCCCTGGCCCACTACAGTAGGCATGGGCCCTTGGAGCGGCTCC	1362
Qy	1261	CCAGCGACCGAGGAGAGCTCCCAACGACGACGAGGCTTCTTCTGACGCCCCCACTCC	1320
Db	1363	CCAGCGACCGAGGAGAGCTCCCAACGACGACGAGGCTTCTTCTGACGCCCCCACTCC	1422
Qy	1321	TCTAAGCCCCGAGGCCCTGTGCTGCGCAGAGCTGAGGAGCGGAGCTTGGCAAG	1380
Db	1423	TCTAAGCCCCGAGGCCCTGTGCTGCGCAGAGCTTGGAGGAGCGGAGCTTGGCAAG	1482
Qy	1381	GCACACTTGAGCGTGCCATCTCTGGGGCAGGAGGCGCTGCGGCTGCACGCCCTGTGC	1440
Db	1483	GCACACTTGAGCGTGCCATCTCTGGGGCAGGAGGCGCTGCGGCTGCACGCCCTGTGC	1542
Qy	1441	CGCGTCTCTGCGGAGGTGACCTGCTTCGGGCTGTGATCTCCCAAGCGCTGACGCTCA	1500
Db	1543	CGCGTCTCTGCGGAGGTGACCTGCTTCGGGCTGTGATCTCCCAAGCGCTGACGCTCA	1602

QY	1501	CTGCCAAAGTATGCGGAGCTCGACCGTGAAGATGACTTCTGTGAGCGTGCAGGCGCCCG	1561
DB	1603	CTGCCAAAGTATGCGGAGCTCGACCGTGAAGATGACTTCTGTGAGCGTGCAGGCGCCCG	1662
QY	1561	GACATCCAGCTTAAGACCCACCAAGAGCCAGAGGCCAGGATGCAACGCTGTGCCAGGGG	1620
DB	1663	GACATCCAGCTTAAGACCCACCAAGAGCCAGAGGCCAGGATGCAACGCTGTGCCAGGGG	1722
QY	1621	AAGGGGCTTGACATCTTCATTCGGCTGGGGCCCTGTCTGTGTCTCAGCCAGAACCGG	1680
DB	1723	AAGGGGCTTGACATCTTCATTCGGCTGGGGCCCTGTCTGTGTCTCAGCCAGAACCGG	1782
QY	1681	TGGCGGCTGTGGGGCCCGCTCCACTGACCCGAGGAGAGGGCGGCTTTGGCCTCACGCTT	1740
DB	1783	TGGCGGCTGTGGGGCCCGCTCCACTGACCCGAGGAGAGGGCGGCTTTGGCCTCACGCTT	1842
QY	1741	CGGGAGACTGCGCTGTCTCATCTGCTGCCGTCAATCCAGGAGCCAGGCGCGGGCGGCT	1800
DB	1843	CGGGAGACTGCGCTGTCTCATCTGCTGCCGTCAATCCAGGAGCCAGGCGCGGGCGGCT	1902
QY	1801	GGCTGAGGAGGGCGACTACATTTGCTCTAGTGAATGGCGAGCCATCCAGTGGTGGAGA	1860
DB	1903	GGCTGAGGAGGGCGACTACATTTGCTCTAGTGAATGGCGAGCCATCCAGTGGTGGAGA	1962
QY	1861	CACGCGAGGTGGTGACGGAGCTGAAGGCTGCGGGAGAGCGGGCGCCAGCTGCAGGTG	1920
DB	1963	CACGCGAGGTGGTGACGGAGCTGAAGGCTGCGGGAGAGCGGGCGCCAGCTGCAGGTG	2022
QY	1921	GTGTGCTGTGCCCGAGCTTAGACTGCCAGCTTTGGGGAGCCGCCGGCCGCTCTGCTG	1980
DB	2023	GTGTGCTGTGCCCGAGCTTAGACTGCCAGCTTTGGGGAGCCGCCGGCCGCTCTGCTG	2082
QY	1981	GGCCCCAGGGGGCTTTAAAGAGCCAGAGGAGCATGTGTTGCAAGACCCCGGCATCCACG	2040
DB	2083	GGCCCCAGGGGGCTTTAAAGAGCCAGAGGAGCATGTGTTGCAAGACCCCGGCATCCACG	2142
QY	2041	TGGGCGAGTCCCGGGCCCTCTCTCACTGAGCCGGAAGGCCCHAGCAGGGCGACACTGGA	2100
DB	2143	TGGGCGAGTCCCGGGCCCTCTCTCACTGAGCCGGAAGGCCCHAGCAGGGCGACACTGGA	2202
QY	2101	GGCTGCCCGCCAGCCCTGTGCCCGCAGTGAAGCCAGCTCGGCCCTCATCTCTTGAAGCACCCA	2160
DB	2203	GGCTGCCCGCCAGCCCTGTGCCCGCAGTGAAGCCAGCTCGGCCCTCATCTCTTGAAGCACCCA	2262
QY	2161	GGGTGGCGG 2169	
DB	2263	GGGTGGCGG 2271	

```

RESULT 2
US-10-274-878-1
; Sequence 1, Application US/10274878
; Patent No. 6670163
; GENERAL INFORMATION:
; APPLICANT: RUSCH, Douglas et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001039-CIP-DIV
; CURRENT APPLICATION NUMBER: US/10/274,878
; CURRENT FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: 09/849,334
; PRIOR FILING DATE: 2001-05-07
; PRIOR APPLICATION NUMBER: 09/773,371
; PRIOR FILING DATE: 2001-02-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2469
; TYPE: DNA
; ORGANISM: Human
US-10-274-878-1

```

Query Match									
Best Local Similarity 100.0%; Score 2169; DB 4; Length 2469;									
Matches 2169; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	ATGATCTCGAGAGAGCGGACGCGCGGGCGCGGAGAGAGCGCGGCTGCAG	60						
DB	103	ATGATCTCGAGAGAGCGGAGCGCGGGCGCGGAGAGAGCGCGGCTGCAG	162						
QY	61	ATATCTAGAGAAAAACCCAGGAAAAACACGCTGTGAGCTCTTTACGGGAAAGACGGGAAGGC	120						
DB	163	ATATCTAGAGAAAAACCCAGGAAAAACACGCTGTGAGCTCTTTACGGGAAAGACGGGAAGGC	222						
QY	121	CTGAGAGACGTGTGCTGTGAGAGAGGCTGTGGGTCCACAGAGGGAAACCCAGTGGCT	180						
DB	223	CTGAGAGACGTGTGCTGTGAGAGAGGCTGTGGGTCCACAGAGGGAAACCCAGTGGCT	282						
QY	181	GTGCACGTTGGCCCAATGAATCCGACGTTTCATGCACTGGGCTGTGACTCCCTGACGCAG	240						
DB	283	GTGCACGTTGGCCCAATGAATCCGACGTTTCATGCACTGGGCTGTGACTCCCTGACGCAG	342						
QY	241	ATCCAGTCGGCCAGTCGAGAGCCGACGGGCCAGATTCAACAGCAGATTGCAAGGAG	300						
DB	343	ATCCAGTCGGCCAGTCGAGAGCCGACGGGCCAGATTCAACAGCAGATTGCAAGGAG	402						
QY	301	CTGCAGATCGGACGGGCGCTGAGAACCTCTACAGAGCCACAGCAACACCGGGTGAG	360						
DB	403	CTGCAGATCGGACGGGCGCTGAGAACCTCTACAGAGCCACAGCAACACCGGGTGAG	462						
QY	361	GAGACGCTCGCCCTGGAGCTGAGCTACGTCACCTCCAACTCGACGCTGCTGAAGAGGAG	420						
DB	463	GAGACGCTCGCCCTGGAGCTGAGCTACGTCACCTCCAACTCGACGCTGCTGAAGAGGAG	522						
QY	421	CTGGAGGAGCTCAGCGGTGGAGCCCTGCGCGCATGGAGCGAACTGTCACTGTC	480						
DB	523	CTGGAGGAGCTCAGCGGTGGAGCCCTGCGCGCATGGAGCGAACTGTCACTGTC	582						
QY	481	CCCATGATCCCTCGGCTGAGAGAGCAAGAGCTGGAATGGTCTACACCGCTGAAG	540						
DB	583	CCCATGATCCCTCGGCTGAGAGAGCAAGAGCTGGAATGGTCTACACCGCTGAAG	642						
QY	541	GAGCTGATCTCAGTGACATTTCTGAGAGGACGGCTCTCTAGAGCGCAAAATCAGGAG	600						
DB	643	GAGCTGATCTCAGTGACATTTCTGAGAGGACGGCTCTCTAGAGCGCAAAATCAGGAG	702						
QY	601	CTGGAGGCTCTGCGGACGCCATGCGGACCCCGACCGGAATGAGTCGGGCTGAGAGCTG	660						
DB	703	CTGGAGGCTCTGCGGACGCCATGCGGACCCCGACCGGAATGAGTCGGGCTGAGAGCTG	762						
QY	661	CTCAGAGCTTATTACACAGCTGTGCTTCTGGAATGGCGCTTCTCAACCTCCAGG	720						
DB	763	CTCAGAGCTTATTACACAGCTGTGCTTCTGGAATGGCGCTTCTCAACCTCCAGG	822						
QY	721	AGCCTCGGGCTTCTTCTCCACTGTGTACGCTACGCTTACTGGGCTCCCGGCCAGCAGCT	780						
DB	823	AGCCTCGGGCTTCTTCTCCACTGTGTACGCTACGCTTACTGGGCTCCCGGCCAGCAGCT	882						
QY	781	GCCCTGGGCTTCTGAGAGGCGACGCTTCTTCTTCAACATCGGTGCCCTCCACACGAGATT	840						
DB	883	GCCCTGGGCTTCTGAGAGGCGACGCTTCTTCTTCAACATCGGTGCCCTCCACACGAGATT	942						
QY	841	GGGGCGCGCAGGACCGCTCTGCAACGAGGTGGCGCGCTTATGAGAGCTTCCAG	900						
DB	943	GGGGCGCGCAGGACCGCTCTGCAACGAGGTGGCGCGCTTATGAGAGCTTCCAG	1002						
QY	901	AGGGCGCTGGGGCTTCTCAGCTCTCTGAGGAGAACTTCTCCATGCGCGCCAGCCAGAC	960						
DB	1003	AGGGCGCTGGGGCTTCTCAGCTCTCTGAGGAGAACTTCTTCCATGCGCGCCAGCCAGAC	1062						
QY	961	ATGAGCGTGGCTCTCTGCGCACTGGAGCAGCTCATGAGCTGAGGCTTCCAG	1020						
DB	1063	ATGAGCGTGGCTCTCTGCGCACTGGAGCAGCTCATGAGCTGAGGCTTCCAG	1122						
QY	1021	GTGTTTGAGGGCTCTCACACCTGCTCCATGCGCCCCCAAGACTGCTGCGCCAGCTG	1080						
DB	1123	GTGTTTGAGGGCTCTCACACCTGCTCCATGCGCCCCCAAGACTGCTGCGCCAGCTG	1182						
QY	1081	CGCTGGCGCAGGAGCGCCCGCAGGTGGCAGCCGAGTACAGGCTAGTCACCGGACATG	1140						
DB	1183	CGCTGGCGCAGGAGCGCCCGCAGGTGGCAGCCGAGTACAGGCTAGTCACCGGACATG	1242						
QY	1141	GCCAGCCACCGTCCACGACTACGCTGCTGCTCTCTGACTGCTGCTGCTGCTGCTGCTG	1200						
DB	1243	GCCAGCCACCGTCCACGACTACGCTGCTGCTCTCTGACTGCTGCTGCTGCTGCTGCTG	1302						
QY	1201	GCCAGTACTTCCGCTCCCTGCGCCACTACGCTAGCCATGCGCCCTCTGCGACGGCTCC	1260						
DB	1303	GCCAGTACTTCCGCTCCCTGCGCCACTACGCTAGCCATGCGCCCTCTGCGACGGCTCC	1362						
QY	1261	CCAGGACCGAGGAGAGCTCCCGACGACGAGCAGGTCTTCTGAGCGCCCGCAGCTCC	1320						
DB	1363	CCAGGACCGAGGAGAGCTCCCGACGACGAGCAGGTCTTCTGAGCGCCCGCAGCTCC	1422						
QY	1321	TCTAAGCCCGCAGGCGCTGTGCTGCGCAGGAGCTGGAGAGCGGAGCTGCGGCTGCG	1380						
DB	1423	TCTAAGCCCGCAGGCGCTGTGCTGCGCAGGAGCTGGAGAGCGGAGCTTGGCAAG	1482						
QY	1381	GCACACCTGAAGCGTGCATCTCTGCGGCGCAGGAGGCGCTGCGGCTGCGCCTGTGC	1440						
DB	1483	GCACACCTGAAGCGTGCATCTCTGCGGCGCAGGAGGCGCTGCGGCTGCGCCTGTGC	1542						
QY	1441	CGGCTCTGCGGAGGTGAGCTTCTGCGGCTGTGATCTCCAGAGCTGAGCGCTCA	1500						
DB	1543	CGGCTCTGCGGAGGTGAGCTTCTGCGGCTGTGATCTCCAGAGCTGAGCGCTCA	1602						
QY	1501	CTGGCCAAATGTCGAGCTCGACCTGAGGAGTACTTCTGAGGCTGCGGAGGCGCG	1560						
DB	1603	CTGGCCAAATGTCGAGCTCGACCTGAGGAGTACTTCTGAGGCTGCGGAGGCGCG	1662						
QY	1561	GACATCCAGCTTAAGACCCACAGAACCCAGAGGCGGAGTCCAGCGCTGTCCAGGGG	1620						
DB	1663	GACATCCAGCTTAAGACCCACAGAACCCAGAGGCGGAGTCCAGCGCTGTCCAGGGG	1722						
QY	1621	AAGGGCTCTGACATCTTCTGATGCGGCTGGGCGCCCTGTCTGTGTTCTCAGCCAAACCG	1680						
DB	1723	AAGGGCTCTGACATCTTCTGATGCGGCTGGGCGCCCTGTCTGTGTTCTCAGCCAAACCG	1782						
QY	1681	TGGCGGCTGGTGGGCGCCGTCCACCTGACCGGAGAGGCGGCTTTGGCTCAGCTT	1740						
DB	1783	TGGCGGCTGGTGGGCGCCGTCCACCTGACCGGAGAGGCGGCTTTGGCTCAGCTT	1842						
QY	1741	CGGGGAGACTCGCTGTCTCATCGCTCCGCTCAATTCAGGGAGCGGCGCGCGGCT	1800						
DB	1843	CGGGGAGACTCGCTGTCTCATCGCTCCGCTCAATTCAGGGAGCGGCGCGCGGCT	1902						
QY	1801	GCCCTGAGAGGAGGCGACTCATTTGTGTGCTGATGGGAGCGCATGAGTGTGTGAGA	1860						
DB	1903	GCCCTGAGAGGAGGCGACTCATTTGTGTGCTGATGGGAGCGCATGAGTGTGTGAGA	1962						
QY	1861	CACGCGAGGTGTGACGAGCTGAAGGCTCGGAGAGGCGGCGGCGGCTGCGAGTG	1920						
DB	1963	CACGCGAGGTGTGACGAGCTGAAGGCTCGGAGAGGCGGCGGCGGCTGCGAGTG	2022						
QY	1921	GTGTGCTGTGCCAGCTCTAGACTGCCAGCTTGGGGGACCGCGGCGCTCTCTGCTG	1980						
DB	2023	GTGTGCTGTGCCAGCTCTAGACTGCCAGCTTGGGGGACCGCGGCGCTCTCTGCTG	2082						
QY	1981	GGCCCGAGGGGCTTCTAAGAGCGCAGGAGCATGTTGCAAGACCCCGGCTCCAG	2040						
DB	2083	GGCCCGAGGGGCTTCTAAGAGCGCAGGAGCATGTTGCAAGACCCCGGCTCCAG	2142						
QY	2041	TGGGCGAGTCCCGGCGCTCTCAACTGGAGCCGAAAGGCGGAGGCGGAGGCTGGA	2100						
DB	2143	TGGGCGAGTCCCGGCGCTCTCAACTGGAGCCGAAAGGCGGAGGCGGAGGCTGGA	2202						
QY	2101	GGTGTCCCGCCAGCGCTGTGCGCCAGTGAAGCAGCTCGCGCTCATCTTGAAGCACCA	2160						

Db 2203 GGCTGCCCCAGCCCTGTGCCCCAGTGAAGCAGCTCCGCCCTCATCTTTGAAGCAACCA 2262
QY 2161 GGGTGGCCG 2169
Db 2263 GGGTGGCCG 2271

RESULT 3
US-09-833-381-1293
; Sequence 1293, Application US/09833381
; Patent No. 6672186
; GENERAL INFORMATION:
; APPLICANT: ROBISON, Keith E.
; TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1293
; LENGTH: 1155
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1155)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1293

Query Match 36.3%; Score 787; DB 4; Length 1155;
Best Local Similarity 99.4%; Pred. No. 4e-156;
Matches 790; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1375 GGCAAGGCACACCTGAAGCGTGCATCTCTGGGCGAGGAGGCGCTGCGCTGCAGGCC 1434
Db 133 GGCAAGGCACACCTGAAGCGTGCATCTCTGGGCGAGGAGGCGCTGCGCTGCAGGCC 192
QY 1435 CTGTGCGCGCTCTGCGCGAGTGGACCTGTCTCGGCTGTGATCTCCCAAGCGCTCAG 1494
Db 193 CTGTGCGCGCTCTGCGCGAGTGGACCTGTCTCGGCTGTGATCTCCCAAGCGCTCAG 252
QY 1495 CGCTCACTGCGCAAGTATGCGAGCTGACCGTGAAGATGATCTCTGAGGCTGCGCAG 1554
Db 253 CCCCCACCGGCAAGTATGCGAGCTGACCGTGAAGATGATCTCTGAGGCTGCGCAG 312
QY 1555 GCCCGGACATCCAGCCTTAAGACCCACAGAACCCAGAGGCCAGGATGCCACGCTGTCC 1614
Db 313 GCCCGGACATCCAGCCTTAAGACCCACAGAACCCAGAGGCCAGGATGCCACGCTGTCC 372
QY 1615 CAGGGGAAGGGGCTGACATCTTCATCGGCTGGGGCCCTGTCTGTGTTCTCAGCAAG 1674
Db 373 CAGGGGAAGGGGCTGACATCTTCATCGGCTGGGGCCCTGTCTGTGTTCTCAGCAAG 432
QY 1675 AACCGTGGCGGCTGTGTGGGCGGCTGCACCTGACCCGAGAGAGGGGCTTGGCTC 1734
Db 433 AACCGTGGCGGCTGTGTGGGCGGCTGCACCTGACCCGAGAGAGGGGCTTGGCTC 492
QY 1735 ACGCTTCGGGAGACTCGCCTGTCTCATCGCTGCCGTCAATTCAGGGAGCCAGCGCGCG 1794
Db 493 ACGCTTCGGGAGACTCGCCTGTCTCATCGCTGCCGTCAATTCAGGGAGCCAGCGCGCG 552
QY 1795 GCGCTGCGCTGAAGAGGGGCGACTACATTTGTGTGAGTGAATGGCGAGCCATGCAAGTGG 1854
Db 553 GCGCTGCGCTGAAGAGGGGCGACTACATTTGTGTGAGTGAATGGCGAGCCATGCAAGTGG 612
QY 1855 TGGAGACACCGGAGGTGTGACCGAGCTGAAAGCTGCGGAGAGGGCGGCCAGCGCTG 1914
Db 613 TGGAGACACCGGAGGTGTGACCGAGCTGAAAGCTGCGGAGAGGGCGGCCAGCGCTG 672
QY 1915 CAGGTGTGTGCTGCTGCCAGCTCTAGACTGCCAGCTTTGGGGGACCGCGCGCCCGTC 1974

Db 673 CAGGTGTGTGCTGCTGCCAGCTCTAGACTGCCAGCTTTGGGAGACCGCGCCCGTC 732
QY 1975 CTGCTGGGCCCCAGGGGGCTTTTAAGAGCCAGAGGAGCATGTTGCAAGACCCCGCA 2034
Db 733 CTGCTGGGCCCCAGGGGGCTTTTAAGAGCCAGAGGAGCATGTTGCAAGACCCCGCA 792
QY 2035 TCCACGTGGGCGAGTCCCGGCCCCCTCAACTGGAGCCAAAGCCAGCAGGCGCAAG 2094
Db 793 TCCACGTGGGCGAGTCCCGGCCCCCTCAACTGGAGCCAAAGCCAGCAGGCGCAAG 852
QY 2095 ACTGAGAGGTGCCCCCAGGCCCTGTGCCCCAGTGAAGCCAGCTCCGCCCTCATCTTGAAG 2154
Db 853 ACTGAGAGGTGCCCCCAGGCCCTGTGCCCCAGTGAAGCCAGCTCCGCCCTCATCTTGAAG 912
QY 2155 CACCCAGGGTGGCCG 2169
Db 913 CACCCAGGGTGGCCG 927

RESULT 4
US-09-849-334-3
; Sequence 3, Application US/09849334
; Patent No. 6500655
; GENERAL INFORMATION:
; APPLICANT: RUSCH, Douglas et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001099-CIP
; CURRENT APPLICATION NUMBER: US/09/849,334
; CURRENT FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 19025
; TYPE: DNA
; ORGANISM: Human
US-09-849-334-3

Query Match 10.0%; Score 217; DB 4; Length 19025;
Best Local Similarity 76.6%; Pred. No. 2e-36;
Matches 321; Conservative 0; Mismatches 10; Indels 88; Gaps 1;

QY 1638 CCATCGGCTGCGGCGCCCTGTCTGTGTTCTCAGCCAAAGACCGGTGGCGGCTGTGTGGGCC 1697
Db 14938 CCGCTGCCAGGGGCGCCCTGTCTGTGTTCTCAGCCAAAGACCGGTGGCGGCTGTGTGGGCC 14997
QY 1698 CGTCCACTGAACCGAGAGAGGCGGCTTTGGCTTCAAGCTTCGGGGAGACTGCGCTGT 1757
Db 14998 CGTCCACTGAACCGAGAGAGGCGGCTTTGGCTTCAAGCTTCGGGGAGACTGCGCTGT 15057
QY 1758 CCTCATGCTGCGCTGATTCAGGGAGCCAGGCGC----- 1793
Db 15058 CCTCATGCTGCGCTGATTCAGGGAGCCAGGCGCGGTGAAGGCGCGCGGCCCCCGCTG 15117
QY 1794 ----- 1793
Db 15118 AGGCTGAGTCTTGGTCCAGCAGGCTGTCTGTCTCCCACTCACCGTCCAGTCTCCC 15177
QY 1794 -----GGCGGCTGGCCTGAAGAGAGGCGCACTACATTTGTGTGATGAGTGGGAGCAGCATGCA 1849
Db 15178 CACAGCGGCTGGCTGAAGAGGCGCACTACATTTGTGTGATGAGTGGGAGCAGCATGCA 15237
QY 1850 GGTGTGGAGACACCGGAGGTGTGACGGAGCTGAAGGCTGCGGGAGAGCGCGGCGCA 1909
Db 15238 GGTGTGGAGACACCGGAGGTGTGACGGAGCTGAAGGCTGCGGGAGAGCGCGGCGCA 15297
QY 1910 GCCTGCAAGGTGTGCTGCTGCCAGCTCTAGACTGCCAGCTTTGGGGAGACCGCGCG 1968
Db 15298 GCCTGCAAGGTGTGCTGCTGCCAGCTCTAGACTGCCAGCTTTGGTGAGCCCGCTGG 15356

```
RESULT 5
US-10-274-878-3
; Sequence 3, Application US/10274878
; Patent No. 6670163
; GENERAL INFORMATION:
; APPLICANT: RUSCH, Douglas et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001099-CIP-DIV
; CURRENT APPLICATION NUMBER: US/10/274,878
; CURRENT FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: 09/849,334
; PRIOR FILING DATE: 2001-05-07
; PRIOR APPLICATION NUMBER: 09/773,371
; PRIOR FILING DATE: 2001-02-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 19025
; TYPE: DNA
; ORGANISM: Human
US-10-274-878-3

Query Match
Best Local Similarity 10.0%; Score 217; DB 4; Length 19025;
Matches 321; Conservative 0; Mismatches 10; Indels 88; Gaps 1;

QY 1638 CCATCGGTGGGGCCCTGTCTGTGTTCTCAGCCAAAGAACCGGTGGCGCTGTGGGGCC 1697
Db 14938 CCCCTGCCAGGGGGCCCTGTCTGTGTTCTCAGCCAAAGAACCGGTGGCGCTGTGGGGCC 14997

QY 1698 GTCACACCTGACCCGAGAGAGAGGGGCTTTGGCTCAGCTTCGGGGAGACTCGCTGT 1757
Db 14998 GTCCACCTGACCCGAGAGAGAGGGGCTTTGGCTCAGCTTCGGGGAGACTCGCTGT 15057

QY 1758 CTTATCGTGGCGTCAATTCAGAGGAGCCAGGCCGCGC----- 1793
Db 15058 CTTATCGTGGCGTCAATTCAGAGGAGCCAGGCCGCGCGTAAGGGCCCGCGGCCCTGT 15117

QY 1794 ----- 1793
Db 15118 AGGCTGAGTCTTGTGTCAGCAGCGGTGCTCTGTCCTCCACCTCACCGTCCAAATCTCC 15177

QY 1794 -----GGCGCTGGCTTGAAGAGGGCGACTACATTTGTGTCAGTGAATGGCAGCATGCA 1849
Db 15178 CACAGGCGGCTGGCTGGAAGAGGGCGACTACATTTGTGTCAGTGAATGGCAGCATGCA 15237

QY 1850 GTGTGTGAGACACACCGAGAGTGGTGACGAGCTGAAGCTCGGGGAGAGGGCGGCGCA 1909
Db 15238 GTGTGTGAGACACACCGAGAGTGGTGACGAGCTGAAGCTCGGGGAGAGGGCGGCGCA 15297

QY 1910 GCCTCAGGTGGTGTGCTGTGCCAGCTCTAGACTGCCAGCTTGGGGGACCGCGG 1968
Db 15298 GCCTCAGGTGGTGTGCTGTGCCAGCTCTAGACTGCCAGCTTGGTGAGCCCTGTG 15356

RESULT 6
US-09-849-334-1/c
; Sequence 1, Application US/09849334
; Patent No. 6500655
; GENERAL INFORMATION:
; APPLICANT: RUSCH, Douglas et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001099-CIP
; CURRENT APPLICATION NUMBER: US/09/849,334
; CURRENT FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2469
; TYPE: DNA
; ORGANISM: Human
US-09-849-334-1
```

```
; TYPE: DNA
; ORGANISM: Human
US-09-849-334-1

Query Match
Best Local Similarity 3.1%; Score 68; DB 4; Length 2469;
Matches 158; Conservative 0; Mismatches 150; Indels 0; Gaps 0;

QY 1344 GCCGAGAGCTGGAGGAGCGCAGCTTGGCAAGGCACACCTGAAGCGTGCCATCCT 1403
Db 1753 GCCCAGCGATGGAGATGTTCAGGCCCTTCCCTGGGACAGCGCTGGCATCTGGCCT 1694

QY 1404 GGGGAGAGGAGGCGCTCGGCTGCACGCCCTGTGTCGCCGCTCTCTCGGAGGTGACCT 1463
Db 1693 CTGGCTTCTGTGGGTCTTAGGCTGGATGTCCGGGCGCTCGCAGCCTCACAGAAGTCAT 1634

QY 1464 GCTTCGGGCTGTGATCTCCACAGCGCTGCAGCGCTCACTGGCCAAAGTATGCGGAGTCGA 1523
Db 1633 CTTCAGGTCGAGCTCCGATACTTTGGCCAGTGGAGCGCTTGGGAGATCACAG 1574

QY 1524 CCGTGAGGATGACTTCTGTGAGGCTGCCGAGGCCCGGACATCCAGCCTAAGACCCACCA 1583
Db 1573 CCGAAGCAGGTCCACCTCGCGAGACGCGGACAGGGCGTGGAGCCCGCCTCCT 1514

QY 1584 GAAGCAGAGGCGCAGATGCCACGCTGTCCAGGGGAGGGGCGCTTGACATCTTCCATCG 1643
Db 1513 CCGGCCAGGATGCGAGCTTCAGGTGTGCTTGCAGGCTGCTTGCAGGCTGCTTCCAGCT 1454

QY 1644 GCTGGGGC 1651
Db 1453 CCGCGGC 1446

RESULT 7
US-10-274-878-1/c
; Sequence 1, Application US/10274878
; Patent No. 6670163
; GENERAL INFORMATION:
; APPLICANT: RUSCH, Douglas et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001099-CIP-DIV
; CURRENT APPLICATION NUMBER: US/10/274,878
; CURRENT FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: 09/849,334
; PRIOR FILING DATE: 2001-05-07
; PRIOR APPLICATION NUMBER: 09/773,371
; PRIOR FILING DATE: 2001-02-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2469
; TYPE: DNA
; ORGANISM: Human
US-10-274-878-1

Query Match
Best Local Similarity 3.1%; Score 68; DB 4; Length 2469;
Matches 158; Conservative 0; Mismatches 150; Indels 0; Gaps 0;

QY 1344 GCCGAGAGCTGGAGGAGCGCAGCTTGGCAAGGCACACCTGAAGCGTGCCATCCT 1403
Db 1753 GCCCAGCGATGGAGATGTTCAGGCCCTTCCCTGGGACAGCGCTGGCATCTGGCCT 1694

QY 1404 GGGGAGAGGAGGCGCTCGGCTGCACGCCCTGTGTCGCCGCTCTCTCGGAGGTGACCT 1463
Db 1693 CTGGCTTCTGTGGGTCTTAGGCTGGATGTCCGGGCGCTCGCAGCCTCACAGAAGTCAT 1634

QY 1464 GCTTCGGGCTGTGATCTCCACAGCGCTGCAGCGCTCACTGGCCAAAGTATGCGGAGTCGA 1523
Db 1633 CTTACAGGTGAGCTCCGATACTTTGGCCAGTGGAGCGCTTGGGAGATCACAG 1574
```


[illegible]

QY 1391 AGCGTGCCATCTCGGGCAGGAGGCGCGTGGGCTGACGCCCTGTGCGGGTCTCTGC 1450
DB 22211 GCGGGGCGGCGGCTCTCGCGGTGCGCTGAGCCAGGCGCCACCTGGACCGGAGGCTC 22270
QY 1451 GCGAGGTGACCTGCTTCGGGCTGTGATCTCCAGACGCTGACGCGCTCACTGGCCCAAGT 1510
DB 22271 TCGCGGAGCACCTGCGCCAGGCTTGCGGAGACCGCGCGCTCGGGGCTGCTCTCGC 22330
QY 1511 ATGCGGAGTCAACGCTGAGGATGACTTCTGTGAGGCTGCGGAGGCGCCCGGACATC 1566
DB 22331 TCCTCGCCCTCGACGAAAGTCCCTCGCGGACCATGCGCGCTGCGCGGGGACTC 22386

RESULT 12

US-08-457-646A-6

; Sequence 6, Application US/08457646A

; Patent No. 5679560

; GENERAL INFORMATION:

; APPLICANT: Schupp, Thomas

; APPLICANT: Ligon, James M.

; APPLICANT: Beck, James Joseph

; APPLICANT: Hill, Dwight Steven

; APPLICANT: Ryals, John Andrew

; APPLICANT: Gaffney, Thomas Deane

; APPLICANT: Lam, Stephen Ting

; APPLICANT: Hammer, Phillip E.

; APPLICANT: Uknes, Scott Joseph

; TITLE OF INVENTION: Genes for the synthesis of

; NUMBER OF SEQUENCES: 22

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Ciba-Geigy Corporation

; STREET: 7 Skyline Drive

; CITY: Hawthorne

; STATE: NY

; COUNTRY: USA

; ZIP: 10532

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/457,646A

; FILING DATE: 01-JUN-1995

; CLASSIFICATION: 530

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/457,205

; FILING DATE: 01-JUN-1995

; APPLICATION NUMBER: 08/258,261

; FILING DATE: 08-Jun-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Elmer, James Scott

; REGISTRATION NUMBER: 36,129

; REFERENCE/DOCKET NUMBER: CGC 1506/CIP3

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 919-541-8614

; TELEFAX: 919-541-8689

; INFORMATION FOR SEQ ID NO: 6:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 28958 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

; US-08-457-646A-6

Query Match

Best Local Similarity 2.6%; Score 57.2; DB 1; Length 28958;

Matches 267; Conservative 0; Mismatches 323; Indels 6; Gaps 1;

QY 971 CBTCCCTCTGCGACTGCGAGCTCATGATGCGCCAGGAGATGTGTTTGGG 1030
DB 21797 CGACGATCTGCCCCCGGGAAGCGGTGCGCTGCCACCTACCCCTCCAGCGGAGC 21856
QY 1031 GCCTCTCACCACTGCTTCATGCGCCCCCAAGACTGCTGCGCCAGCTGCGCGC 1090
DB 21857 GTTCTGGCTCGACGCTTCACGCGACCGCGCGCGGCTCAACCACTTGTCTCGCTCG 21916
QY 1091 AGGAGCGCCCGCAGGTGGCAGCCGAGTACAGAGCTAGTGCACCGGACCATGCGCCGAC 1150
DB 21917 AGGGGGGTTCGCGAGCCCATCGAGCGGAATATGACCGGTTCAGCGCCAGCTCC 21976
QY 1151 CCGTCCAGACTACCTGCTCTCTGACTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1210
DB 21977 AGCTGAGCGGCGAGCAGCGCGCGCTTGGCTGCTGCTTCCACCTCGCGAGCT 22036
QY 1211 TCCGCTCCCTGCGCCACTACAGCTAGCATGCGCTCTGCGACGCGCTCCCGAGGACCG 1270
DB 22037 TTCGCCAGCGCGCAAGAGCAGGCGACGCTGCGCGCTACCGCATCACTGGA 22096
QY 1271 AGGAGAGCTCCCGCAGCAGCAGGAGTCTTCTGCGAGCGCGCGCGCTGCTGCTGCTGCT 1330
DB 22097 AGCTCTGACACCGCGCAGCGCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 22150
QY 1331 GAGGCCCTGCTGCTGCGCGAGGAGCTGAGGAGCGCAGGAGCTTGGCAAGGCACACCTGA 1390
DB 22151 TCGTGGCGCGCTCTGAGCAGCAGCGCTCCCTCGCGCTCACCGAGGCGCTCGCCC 22210
QY 1391 AGCGTGCCATCTGCGGCGAGGAGGCGCTGCGGCTGCGCGCTGCGCGCTGCGCGCTGCG 1450
DB 22211 GCGCGCGCGCGCTGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCT 22270
QY 1451 GCGAGGTGACCTGCTTCGGGCTGTGATCTCCAGACGCTGCGAGCGCTCACCTGGCCAAGT 1510
DB 22271 TCGCGGAGCAGCTGCGCGCGCTTGGCGCGAGACCGCGCGCTGCGCGCGCTGCGCGCT 22330
QY 1511 ATGCGGAGCTCGACCGTGGAGTGAATCTTCTGTGAGGCTGCGGAGCGCCCGGACATC 1566
DB 22331 TCCTCGCCCTCGAGAAAGTCCCTCGCGGACCATGCGCGCTGCGCGCGCTGCGCGCTGCG 22386

RESULT 13

US-08-458-076A-6

; Sequence 6, Application US/08458076A

; Patent No. 5698425

; GENERAL INFORMATION:

; APPLICANT: Schupp, Thomas

; APPLICANT: Ligon, James M.

; APPLICANT: Beck, James Joseph

; APPLICANT: Hill, Dwight Steven

; APPLICANT: Ryals, John Andrew

; APPLICANT: Gaffney, Thomas Deane

; APPLICANT: Lam, Stephen Ting

; APPLICANT: Hammer, Phillip E.

; APPLICANT: Uknes, Scott Joseph

; TITLE OF INVENTION: Genes for the synthesis of

; NUMBER OF SEQUENCES: 22

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Ciba-Geigy Corporation

; STREET: 7 Skyline Drive

; CITY: Hawthorne

; STATE: NY

; COUNTRY: USA

; ZIP: 10532

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/458,076A

; FILING DATE: 01-JUN-1995

```

; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/457,205
; FILING DATE: 01-JUN-1995
; APPLICATION NUMBER: 08/258,261
; FILING DATE: 08-Jun-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Elmer, James Scott
; REGISTRATION NUMBER: 36,129
; REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8614
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28958 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-458-076A-6

Query Match 2.6%; Score 57.2; DB 1; Length 28958;
Best Local Similarity 44.8%; Pred. No. 0.0067;
Matches 267; Conservative 0; Mismatches 323; Indels 6; Gaps 1;

QY 971 CGTCCCTCTCGCAGCTGAGCAGCTCATGATGGCCAGGCCCCAGGAGATGTGTTTGAGG 1030
Db 21797 CGAGATCTCTGCCCCCGGGAAGCGGTGCGCTGCCACCTACCCCTTCAGCGCGAGC 21856

QY 1031 GCCTCTCACACCTGCTCCATGCGCCCCCAAGACTGCTGCGCCAGCTGGCCAGCTGGCTGGCGC 1090
Db 21857 GCTTCTGGCTCGAGCGCTCCACGSCACCCGCGCGCGTCAACACCTTGTCTCGCTCG 21916

QY 1091 AGGAGGCGCGCCAGGTGGCAGCCGAGTACAGGCTAGTGACCGGACCATGCCCCAGCCAC 1150
Db 21917 AGGGCGGTCTTGGCAGGCCATCGAGAGCGGGAATATCGAGCGCTCAGCGGCCAGCTCC 21976

QY 1151 CCGTCCACGACTACGTGCTCTTCTCTGGAAGTGTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 1210
Db 21977 ACGTGGACGCGCAGCAGCAGCGCGCGCCCTTGCCTGCTGCTTCCACCCCTCCGAGCT 22036

QY 1211 TCAGCTCCCTGGCCCACTACACGTAGCATGCGCTCTGCGAGCGCTCCCGAGCAGC 1270
Db 22037 TTCGCCACGAGCGGCAAGACGAGCAGGCGCAGGTGCGAGCGCTGCGCGTACCGCATCACGTGA 22096

QY 1271 AGGAGAGCTCCCCACGACGAGCAGGTCTTCTGACGCGCCCGCCACCTCTCTAAGCCCC 1330
Db 22097 AGCTCTGACACCGCCACACGCGCCCGGACCTGG-----CGGGACCTGGCTCTCG 22150

QY 1331 GAGGCCCTGTGCTGCGCAGGAGCTGAGGAGCGCAGCGAGCTTGGCAAGGCAACCTGA 1390
Db 22151 TCGTGGCGCGCTCTGAGCAGCAGCGCTCCCTCCGCGCTCAGCGAGCGCTCGCCC 22210

QY 1391 AGCGTGCACTCTGGGCGAGAGAGCGGTGCGGCTGCGCTGCGCGCTGCGCGCTCTGCG 1450
Db 22211 GCGCGCGCGCGCGCTGCTGCGCGCTGCGCTGCGCGCGCGCTGCGCGCGCGCTGCGCG 22270

QY 1451 GCGAGGTGGACCTGCTTGGGCTGTGATCTCCAGAGCGCTGACGCGCTCACTGGCCAAAGT 1510
Db 22271 TCGCGGAGCCTTGGCGAGGCTTGGCGCGAGACCGGCGCGCTCGCGGCGGTGCTCTCG 22330

QY 1511 ATGCGGAGCTCGACGCTGAGGATGACTTCTGTGAGGCTGCGAGGCGCCCGGACATC 1566
Db 22331 TCCTCGCCCTCGAGCAAGTCCCTCGCGGACCATGCGCGCTGCGCGCGGCGGACTC 22386

RESULT 14
US-08-764-233A-4
; Sequence 4, Application US/08764233A
; Patent No. 5716849

; GENERAL INFORMATION:
; APPLICANT: Ligon, James M.
; APPLICANT: Schupp, Thomas
; APPLICANT: Beck, James J.
; APPLICANT: Hill, Dwight S.
; APPLICANT: Neff, Snezana
; APPLICANT: Ryals, John A.
; TITLE OF INVENTION: Genes For The Biosynthesis Of Soraphen
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ciba-Geigy Corporation
; STREET: 520 White Plains Road, P.O. Box 2005
; CITY: Tarrytown
; STATE: NY
; COUNTRY: USA
; ZIP: 10591
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/764,233A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/729,214
; FILING DATE: 09-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/258,261
; FILING DATE: 08-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: 1506/CIP6
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919) 541-8587
; TELEFAX: (919) 541-8689
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28958 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Sorangium cellulosum
; IMMEDIATE SOURCE:
; CLONE: p98/1
; US-08-764-233A-4

Query Match 2.6%; Score 57.2; DB 1; Length 28958;
Best Local Similarity 44.8%; Pred. No. 0.0067;
Matches 267; Conservative 0; Mismatches 323; Indels 6; Gaps 1;

QY 971 CGTCCCTCTCGCAGCTGAGCAGCTCATGATGGCCAGGCCCCAGGAGATGTGTTTGAGG 1030
Db 21797 CGAGATCTCTGCCCCCGGGAAGCGGTGCGCTGCCACCTACCCCTTCAGCGCGAGC 21856

QY 1031 GCCTCTCACACCTGCTCCATGCGCCCCCAAGACTGCTGCGCCAGCTGGCCAGCTGGCTGGCGC 1090
Db 21857 GCTTCTGGCTCGAGCGCTCCACGSCACCCGCGCGCGTCAACACCTTGTCTCGCTCG 21916

QY 1091 AGGAGGCGCGCCAGGTGGCAGCCGAGTACAGGCTAGTGACCGGACCATGCCCCAGCCAC 1150
Db 21917 AGGGCGGTCTTGGCAGGCCATCGAGAGCGGGAATATCGAGCGCTCAGCGGCCAGCTCC 21976

QY 1151 CCGTCCACGACTACGTGCTCTTCTCTGGAAGTGTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 1210
Db 21977 ACGTGGACGCGCAGCAGCAGCGCGCGCCCTTGCCTGCTGCTTCCACCCCTCCGAGCT 22036

QY 1211 TCAGCTCCCTGGCCCACTACACGTAGCATGCGCTCTGCGAGCGCTCCCGAGCAGC 1270
Db 22037 TTCGCCACGAGCGGCAAGACGAGCAGGCGCAGGTGCGAGCGCTGCGCGTACCGCATCACGTGA 22096

QY 1271 AGGAGAGCTCCCCACGACGAGCAGGTCTTCTGACGCGCCCGCCACCTCTCTAAGCCCC 1330
Db 22097 AGCTCTGACACCGCCACACGCGCCCGGACCTGG-----CGGGACCTGGCTCTCG 22150

QY 1331 GAGGCCCTGTGCTGCGCAGGAGCTGAGGAGCGCAGCGAGCTTGGCAAGGCAACCTGA 1390
Db 22151 TCGTGGCGCGCTCTGAGCAGCAGCGCTCCCTCCGCGCTCAGCGAGCGCTCGCCC 22210

QY 1391 AGCGTGCACTCTGGGCGAGAGAGCGGTGCGGCTGCGCTGCGCGCTGCGCGCTCTGCG 1450
Db 22211 GCGCGCGCGCGCGCTGCTGCGCGCTGCGCTGCGCGCGCGCTGCGCGCGCGCTGCGCG 22270

QY 1451 GCGAGGTGGACCTGCTTGGGCTGTGATCTCCAGAGCGCTGACGCGCTCACTGGCCAAAGT 1510
Db 22271 TCGCGGAGCCTTGGCGAGGCTTGGCGCGAGACCGGCGCGCTCGCGGCGGTGCTCTCG 22330

QY 1511 ATGCGGAGCTCGACGCTGAGGATGACTTCTGTGAGGCTGCGAGGCGCCCGGACATC 1566
Db 22331 TCCTCGCCCTCGAGCAAGTCCCTCGCGGACCATGCGCGCTGCGCGCGGCGGACTC 22386

RESULT 14
US-08-764-233A-4
; Sequence 4, Application US/08764233A
; Patent No. 5716849
```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 12, 2004, 19:58:24 ; Search time 559.787 Seconds
(without alignments)
16460.437 Million cell updates/sec

Title: US-10-697-266-1_COPY_103_2271

Perfect score: 2169
Sequence: 1 atgactctggaggagagcc.....tgaagcaccacgggtggccg 2169

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq 29Jan04:*

1: Geneseqn1980s: *
2: Geneseqn1990s: *
3: Geneseqn2000s: *
4: Geneseqn2001as: *
5: Geneseqn2001bs: *
6: Geneseqn2002s: *
7: Geneseqn2003as: *
8: Geneseqn2003bs: *
9: Geneseqn2003cs: *
10: Geneseqn2004s: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2169	100.0	2469	6	AAL48837 Human pro
2	1951	89.9	2837	7	ABV72511 Nucleotid
3	798.8	36.8	3213	7	ACC46698 Human dit
4	781.4	36.0	1013	5	AAS00831 Human CDN
5	779.8	36.0	1013	3	AAC77780 Human can
6	496.4	22.9	2100	5	AAS15828 Human cDN
7	496.4	22.9	3526	6	AAS15827 Human ORF
8	494.8	22.8	2109	6	ABX97183 Human NOV
9	493.2	22.7	2061	6	ABX71856 Human GTP
10	493.2	22.7	3484	6	ABX71855 Human GTP
11	487.8	22.5	3019	4	AAF58361 Human GTP
12	463.6	21.4	2310	6	ABX97182 Human NOV
13	436	20.1	764	2	AAZ15546 Human gen
14	418.2	19.3	2757	4	AAC91349 Human pol
15	379.2	17.5	1671	6	ABK92255 Prostate
16	355	16.4	3335	5	AAS72556 DNA encod
17	301.4	13.9	2869	4	AAH75576 Human tra
18	235.4	10.9	966	5	AAS70406 DNA encod
19	219.4	10.1	331	3	AAC76200 Human ORF
20	217	10.0	19025	6	AAL48838 Human pro
21	206.6	9.5	2168	4	ABL14125 Drosophil
22	203.8	9.4	735	3	AAC75784 Human ORF
23	165	7.6	165	2	AAZ14379 Human gen

24	150	6.9	933	5	AAS72551	Aas72551 DNA encod
25	143.8	6.6	1330	5	AAS70407	Aas70407 DNA encod
26	108.4	5.0	2925	7	ADA52688	Ada52688 Human cod
27	108.4	5.0	5200	6	AAD40740	Aad40740 Human kin
28	108.2	5.0	5234	3	AAC81224	Aac81224 Human his
29	103.6	4.8	599	4	AAL00243	Aal00243 Human rep
30	97	4.5	802	4	AAS03028	Aas03028 Human dia
31	94.6	4.4	439	8	ACH17102	Ach17102 Human adu
32	93.4	4.3	338	4	AAS57421	Aas57421 cDNA #97
33	87	4.0	479	6	ABL81142	Ab181142 Human ova
34	72.4	3.3	500	6	ABS71881	Ab571881 Human GTP
35	71	3.3	167	5	AAS15837	Aas15837 Human pro
36	71	3.3	389	5	AAS72552	Aas72552 DNA encod
37	69.4	3.2	167	6	ABS71866	Ab571866 Human GTP
38	68.2	3.1	789	5	AAS72550	Aas72550 DNA encod
C 39	68	3.1	2469	6	AAL48837	Aal48837 Human pro
C 40	68	3.1	2837	7	ABV72511	Abv72511 Nucleotid
C 41	66.2	3.1	492	3	AAC74860	Aac74860 Human ORF
C 42	63.6	2.9	2000	7	ADA71938	Ada71938 Rice gene
C 43	63.2	2.9	114955	2	AAX53491	Aax53491 Human ade
C 44	62.4	2.9	507	4	AAL34939	Aal34939 Human mus
C 45	62.4	2.9	507	7	ABX57927	Abx57927 cDNA enco

ALIGNMENTS

RESULT 1

AAL48837
ID AAL48837 standard; cDNA; 2469 BP.

AC AAL48837;

DT 24-OCT-2002 (first entry)

DE Human protein kinase N family kinase coding sequence.

KW Human; protein kinase N; kinase; PKN; eye retinoblastoma; bocio tumour;
KW placenta choriocarcinoma; germ cell; leukocyte; uterus tumour; cancer;
KW pre-B cell acute lymphoblastic leukaemia; Wilm's tumour; enzyme;
KW brain anaplastic oligodendroma; uterus endometrial adenocarcinoma;
KW cytostatic; gene therapy; chromosome 8; gene; ss.

OS Homo sapiens.

PH Key Location/Qualifiers

FT 5'UTR 1..102

FT CDS /*tag= a

FT CDS 103..2274

FT CDS /*tag= b

FT 3'UTR /product= "kinase"

FT CDS 2275..2469

FT CDS /*tag= c

WO200261062-A2.

08-AUG-2002.

29-JAN-2002; 2002WO-US002152.

01-FEB-2001; 2001US-00773371.

07-MAY-2001; 2001US-00849334.

(PEKE) PE CORP NY.

Rusch D, Ketchum KA, Di Francesco V, Beasley EM;

WPI; 2002-608516/65.

P-PSDB; AAO18602.

PT New human kinase peptide and nucleic acid molecule, useful for treating disorders associated with abnormal expression of kinase protein, e.g. retinoblastoma, Wilm's tumor, in drug screening assays and

PT pharmacogenomic analysis.

Claim 4; Fig 1; 76pp; English.

CC The present invention provides the protein and coding sequences of a
CC novel human protein kinase N family protein. The sequences can be used in
CC the treatment of disorders associated with the absence of, inappropriate,
CC or unwanted expression of the protein, e.g. eye retinoblastoma, Wilms
CC tumour, placenta choriocarcinomas, bocio tumours, pre-B cell acute
CC lymphoblastic leukaemias, uterus tumours, brain anaplastic oligodendromas
CC and uterine endometrial adenocarcinomas. The present sequence is the cDNA
CC of the invention. The gene is found on human chromosome 8

XX Sequence 2469 BP; 461 A; 812 C; 821 G; 375 T; 0 U; 0 Other;

Query Match 100.0%; Score 2169; DB 6; Length 2469;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2169; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGATCTGGAGAGAGCGGACGGCGGGGCGCGGAGAGAGAGCGCGGGCTGCAG 60

DB 103 ATGATCTGGAGAGAGCGGACGGCGGGGCGCGGAGAGAGAGCGCGGGCTGCAG 162

QY 61 ATATCTAGGAGAAACCCAGGAAACACAGTGTGAGCTCTTTACGGGAAACAGCGAAGGC 120

DB 163 ATATCTAGGAGAAACCCAGGAAACACAGTGTGAGCTCTTTACGGGAAACAGCGAAGGC 222

QY 121 CTGAGAGACGTGTGTGGTGGAGAGGTGTGGGTCCACAGAGGGGAAGACCCAGTGCCT 180

DB 223 CTGAGAGACGTGTGTGGTGGAGAGGTGTGGGTCCACAGAGGGGAAGACCCAGTGCCT 282

QY 181 GTGACGTTGGCCCATGAATTCGACGCTTCATGACGTGGGCTGTGACTCCCTGACGCGAG 240

DB 283 GTGACGTTGGCCCATGAATTCGACGCTTCATGACGTGGGCTGTGACTCCCTGACGCGAG 342

QY 241 ATTCAGTGGCGGACGTGCAGAGCGGAGGCGGAGGCGGAGATTCACAGCAGATTGACAAGGAG 300

DB 343 ATTCAGTGGCGGACGTGCAGAGCGGAGGCGGAGGCGGAGATTCACAGCAGATTGACAAGGAG 402

QY 301 CTGAGATCGGAGCGGCGGTGAGAACCTCTACAGAGCCACAGCAACACCGGGTGAGA 360

DB 403 CTGAGATCGGAGCGGCGGTGAGAACCTCTACAGAGCCACAGCAACACCGGGTGAGA 462

QY 361 GAGACGTTCCCTTGGAGCTGAGTACGTCAATCCAACTGACGTGCTGAGGAGGAG 420

DB 463 GAGACGTTCCCTTGGAGCTGAGTACGTCAATCCAACTGACGTGCTGAGGAGGAG 522

QY 421 CTGAGAGCTCAGCGGTGGCGTGGACCCCTGGCGGATGGGAGCGAAGCTGTCACTGTC 480

DB 523 CTGAGAGCTCAGCGGTGGCGTGGACCCCTGGCGGATGGGAGCGAAGCTGTCACTGTC 582

QY 481 CCCATGATCCCTTGGCGCTGAAGGAGACCAAGAGCTGAGTGTCTACACCGCTGAAG 540

DB 583 CCCATGATCCCTTGGCGCTGAAGGAGACCAAGAGCTGAGTGTCTACACCGCTGAAG 642

QY 541 GAGCTGATCTCAGTGCATTTGGAGAGACCGCGCTCTACAGGAGCAAGATCAGGAG 600

DB 643 GAGCTGATCTCAGTGCATTTGGAGAGACCGCGCTCTACAGGAGCAAGATCAGGAG 702

QY 601 CTGAGAGCCCTTGGCGGACGATCGGACCCCGGAGCGGAATGAGTCGGGCTGGAGCTG 660

DB 703 CTGAGAGCCCTTGGCGGACGATCGGACCCCGGAGCGGAATGAGTCGGGCTGGAGCTG 762

QY 661 CTCACAGCTATTACACACAGCTGTCTCTCTGATGCGCGCTCTCCTACCCCTGCCAGG 720

DB 763 CTCACAGCTATTACACACAGCTGTCTCTCTGATGCGCGCTCTCCTACCCCTGCCAGG 822

QY 721 AGCTCTGGGCTCTTCTTCCACTGGTACGACTCGCTTACTGGGGTCCCGGCGCCAGCAGCGT 780

DB 823 AGCTCTGGGCTCTTCTTCCACTGGTACGACTCGCTTACTGGGGTCCCGGCGCCAGCAGCGT 882

QY 781 GCCCTGGCTTCGAGAAAGGAGCGGCTTCTTCTTCAACATCGGTGGCTTCCACAGCAGATT 840

DB 883 GCCCTGGCTTCGAGAAAGGAGCGGCTTCTTCTTCAACATCGGTGGCTTCCACAGCAGATT 942

QY 841 GGGGCGGCGCAGAGACCGCTCTGACCCGAGGGTGGCCGCGCTATGAGGCGCTTCCAG 900

DB 943 GGGGCGGCGCAGAGACCGCTCTGACCCGAGGGTGGCCGCGCTATGAGGCGCTTCCAG 1002

QY 901 AGGGCGGCTGGGCGCTTTCAGCCCTCTGAGGGAGAACTTCTCCCATGCGCGAGCCAGAC 960

DB 1003 AGGGCGGCTGGGCGCTTTCAGCCCTCTGAGGGAGAACTTCTCCCATGCGCGAGCCAGAC 1062

QY 961 ATGAGCGCTGCGTCTCTCTGCGACTGAGAGCAGCTCATGATGGCCAGGCCAGGATGT 1020

DB 1063 ATGAGCGCTGCGTCTCTCTGCGACTGAGAGCAGCTCATGATGGCCAGGCCAGGATGT 1122

QY 1021 GTGTTTGGGGCTCTCACCACTGCTCTCATGGCCCCCAAGACTGCCTTGGCCCCAGCTG 1080

DB 1123 GTGTTTGGGGCTCTCTCACCACTGCTCTCATGGCCCCCAAGACTGCCTTGGCCCCAGCTG 1182

QY 1081 CGCTCTGGCGCAGAGAGCGCCCAAGTGGCAGCGAGTACAGGTAGTGACACCGAGCCATG 1140

DB 1183 CGCTCTGGCGCAGAGAGCGCCCAAGTGGCAGCGAGTACAGGTAGTGACACCGAGCCATG 1242

QY 1141 GCCCAGCCACCCCTCCACAGACTAGCTGCTCTCTCTGAGCTGCTGCTGCTGCTGCTGCTG 1200

DB 1243 GCCCAGCCACCCCTCCACAGACTAGCTGCTCTCTCTGAGCTGCTGCTGCTGCTGCTGCTG 1302

QY 1201 GCCCAGTACTTCCGCTCCCTGGGCCCACTACCACTAGCCATGAGCCCTCTGCGAGCGCTCC 1260

DB 1303 GCCCAGTACTTCCGCTCCCTGGGCCCACTACCACTAGCCATGAGCCCTCTGCGAGCGCTCC 1362

QY 1261 CCAGCAGCCAGGAGAGAGTCCCCACGACGACGAGTCTTCTGAGCCCCCAGCCCTCC 1320

DB 1363 CCAGCAGCCAGGAGAGTCCCCACGACGACGAGTCTTCTGAGCCCCCAGCCCTCC 1422

QY 1321 TCTTAAGCCCCGAGGCGCTGTGCTGCGCAGAGCTGAGAGAGCGCAGGAGCTTGGCAAG 1380

DB 1423 TCTTAAGCCCCGAGGCGCTGTGCTGCGCAGAGCTGAGAGAGCGCAGGAGCTTGGCAAG 1482

QY 1381 GCACACTGAAGCGTGCATCTCTGGGCGCAGGAGGCGCTGGGGCTGCGACGCCCTGTGC 1440

DB 1483 GCACACTGAAGCGTGCATCTCTGGGCGCAGGAGGCGCTGGGGCTGCGACGCCCTGTGC 1542

QY 1441 CGGCTCTCGCGCAGAGTGGACCTGCTTTCGGGCTGTGATCTCTCCAGAGCGCTGCGAGCGTCA 1500

DB 1543 CGGCTCTCGCGCAGAGTGGACCTGCTTTCGGGCTGTGATCTCTCCAGAGCGCTGCGAGCGTCA 1602

QY 1501 CTGGCCAGATGTCGGGAGCTGACCGGTGAGGATGACTTCTGTGAGGCTGCGGAGGCCCG 1560

DB 1603 CTGGCCAGATGTCGGGAGCTGACCGGTGAGGATGACTTCTGTGAGGCTGCGGAGGCCCG 1662

QY 1561 GACATCCAGCCTTAAGACCCACAGAGCCAGAGGAGGAGTGCACCGCTGTCCCGAGGGG 1620

DB 1663 GACATCCAGCCTTAAGACCCACAGAGCCAGAGGAGGAGTGCACCGCTGTCCCGAGGGG 1722

QY 1621 AAGGGGCTGACATCTTTCATTCGGCTGGGGCCCCCTGTCTGTGTTCTCAGCCAGAACCGG 1680

DB 1723 AAGGGGCTGACATCTTTCATTCGGCTGGGGCCCCCTGTCTGTGTTCTCAGCCAGAACCGG 1782

QY 1681 TGGCGGCTGTTGGGGCGGCTCCACTGACCCGAGGAGGCGGCTTGGGCTCAGCGCTT 1740

DB 1783 TGGCGGCTGTTGGGGCGGCTCCACTGACCCGAGGAGGCGGCTTGGGCTCAGCGCTT 1842

QY 1741 CGGGGAGACTCGCTCTCTCATCGCTGCCCTCATATCCAGGAGGAGGCGGCGCGGCT 1800

DB 1843 CGGGGAGACTCGCTCTCTCATCGCTGCCCTCATATCCAGGAGGAGGCGGCGCGGCT 1902

QY 1801 GGCCTGAAGAGGCGGAGTACATCTCTCAGTGAATGAGGAGCCAGCTGCTGAGAGA 1860

DB 1903 GGCCTGAAGAGGCGGAGTACATCTCTCAGTGAATGAGGAGCCAGCTGCTGAGAGA 1962

QY 1861 CACCGGAGGCTGCTGAGCGAGCTGAAGGCTCGCGGAGAGGCGGCGCGCAGCTGCTGAGGTG 1920

DB 1963 CACCGGAGGCTGCTGAGCGAGCTGAAGGCTCGCGGAGAGGCGGCGCGCAGCTGCTGAGGTG 2022

QY 1921 GTGTCGCTGTCGCCAGCTCTAGACTGCCAGCTTGGGGAGCCGCGGCCCGTCTGCTG 1980
 Db 2023 GTGTCGCTGTCGCCAGCTCTAGACTGCCAGCTTGGGGAGCCGCGGCCCGTCTGCTG 2082
 QY 1981 GGGCCAGGGGGCTTCTAAGGAGCCAGAGGAGCATGTTGCAAGACCCCGGCATCCAG 2040
 Db 2083 GGGCCAGGGGGCTTCTAAGGAGCCAGAGGAGCATGTTGCAAGACCCCGGCATCCAG 2142
 QY 2041 TGGGCGAGTCCCGGCGCTCTCTCAACTGAGCCGAAAGGCCAGAGGGCAAGACTGGA 2100
 Db 2143 TGGGCGAGTCCCGGCGCTCTCTCAACTGAGCCGAAAGGCCAGAGGGCAAGACTGGA 2202
 QY 2101 GGTCGCCCCAGCCCTGTGCCCCAGTGAAGCAGCTCGCCCTCATCTTGAAGACCCA 2160
 Db 2203 GGTCGCCCCAGCCCTGTGCCCCAGTGAAGCAGCTCGCCCTCATCTTGAAGACCCA 2262
 QY 2161 GGGTGGCCG 2169
 Db 2263 GGGTGGCCG 2271

RESULT 2
 ID ABV72511
 AC ABV72511 standard; cDNA; 2837 BP.
 XX ABV72511;
 DT 29-JAN-2003 (first entry)
 DE Nucleotide sequence of a human intracellular signalling molecule.
 KW Human; intracellular signalling molecule; INTSIG; Alzheimer's disease;
 KW neurodegenerative disorder; Parkinson's disease; muscular disorder;
 KW myotonic dystrophy; catatonia; endocrine disorder; diabetes;
 KW Grave's disease; cancer; leukemia; cervical cancer; breast cancer;
 KW immunological disorder; scleroderma; systemic lupus erythematosus;
 KW allergy; gastrointestinal disorder; Crohn's disease; renal disorder;
 KW Goodpasture's syndrome; infection; cardiovascular disorder;
 KW atherosclerosis; hepatic disease; cirrhosis; gene; ss.
 XX
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FT 122..2134
 CDS /*tag= a
 FT /product= "intracellular signalling molecule"
 XX
 FN W020027235-A2.
 XX
 PD 03-OCT-2002.
 XX
 PF 19-DEC-2001; 2001WO-US050315.
 XX
 XX 21-DEC-2000; 2000US-0257804P.
 PR 05-JAN-2001; 2001US-0260102P.
 XX
 XX (INCY-) INCYTE GENOMICS INC.
 PA
 XX Burford N, Ding L, Yue H, Thornton M, Wallia NK, Gandhi AR;
 PI Arvizu C, Baughn MR, Swarnakar A, Duggan BM, Lu DAM, Thangavelu K;
 PI Warren BA, Tang YT, Khan PA, Yao MG, Emerling BM;
 XX
 DR WPI; 2003-018929/01.
 DR P-PSDB; ABB99400.
 XX
 XX INTSIG polypeptides, useful for diagnosing, preventing, and treating
 PT disorders associated with an abnormal expression or activity of INTSIG,
 PT e.g. neuromuscular, immunological, cardiovascular disorders, cancer and
 PT infection.
 XX
 PS Claim 5; Page 119-120; 126pp; English.
 XX

CC The present sequence encodes a human intracellular signalling molecule.
 CC Such molecules are designated INTSIG. The INTSIG polypeptide or its
 CC fragments, and the INTSIG polynucleotide are useful in diagnosing,
 CC preventing, and treating disorders associated with an abnormal expression
 CC or activity of INTSIG, such as neurodegenerative disorders (e.g.
 CC Parkinson's disease, Alzheimer's disease), muscular disorders (e.g.
 CC myotonic dystrophy, catatonia), endocrine disorders (e.g. diabetes,
 CC Grave's disease), cancers (e.g. leukemia, cervical or breast cancers),
 CC immunological disorders (e.g. scleroderma, systemic lupus erythematosus,
 CC allergies), gastrointestinal disorders (e.g. Crohn's disease), renal
 CC disorders (e.g. Goodpasture's syndrome), infections (e.g. viral,
 CC bacterial, fungal, parasitic, protozoal, helminthic), cardiovascular
 CC disorders (e.g. atherosclerosis), or hepatic diseases (e.g. cirrhosis).
 CC INTSIG or its fragments may also be used in screening for compounds that
 CC specifically bind to and modulate the activity of INTSIG. The
 CC polynucleotides can be used to create humanized animals or transgenic
 CC animals to model human disease
 CC
 XX
 SQ Sequence 2837 BP; 473 A; 993 C; 908 G; 463 T; 0 U; 0 Other;

Query Match 89.9%; Score 1951; DB 7; Length 2837;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1951; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 219 GGGCTGTGACTCCCTGAGCGCAGATCCAGTGCAGCGCCAGCTGCAGAGCCGAGGCCAGAT 278
 Db 181 GGGCTGTGACTCCCTGAGCGCAGATCCAGTGCAGCGCCAGCTGCAGAGCCGAGGCCAGAT 240
 QY 279 TCACGACGAGATTGACAAGAGAGCTGCAGATCGGACGGCGCTGAGAACCTCTACAGAGC 338
 Db 241 TCACGACGAGATTGACAAGAGAGCTGCAGATCGGACGGCGCTGAGAACCTCTACAGAGC 300
 QY 339 CACCAAGCAACACCGGGTGAGAGACCGTGCAGCGCCAGCTGCAGAGCTGAGCTCACTCAAC 398
 Db 301 CACCAAGCAACACCGGGTGAGAGACCGTGCAGCGCCAGCTGCAGAGCTGAGCTCACTCAAC 360
 QY 399 CTTGACGCTGCTGAAAGGAGGAGCTGCAGAGAGCTCAGCGGTGGCGTGGAGCCCTGGCGGCA 458
 Db 361 CTTGACGCTGCTGAAAGGAGGAGCTGCAGAGAGCTCAGCGGTGGCGTGGAGCCCTGGCGGCA 420
 QY 459 TGGGAGCAAGCTGTCACTGCCCATGATCCCCCTGGGCTGAAAGGAGACCAAGAGCT 518
 Db 421 TGGGAGCAAGCTGTCACTGCCCATGATCCCCCTGGGCTGAAAGGAGACCAAGAGCT 480
 QY 519 GACTGTGTACACCGCTGAAGGAGCTGATCTCAGTGACATTTGAGAGGAGCGGCGCTC 578
 Db 481 GACTGTGTACACCGCTGAAGGAGCTGATCTCAGTGACATTTGAGAGGAGCGGCGCTC 540
 QY 579 CTACGAGGCAGAAATCAGGAGCTGGAGGCCCTTGGCGAGGCCATGCGGACCCCCAGCGC 638
 Db 541 CTACGAGGCAGAAATCAGGAGCTGGAGGCCCTTGGCGAGGCCATGCGGACCCCCAGCGC 600
 QY 639 GAATGAGTCGGGCTTGGAGCTGCTCAAGCTATTATCAACAGCTGTCTTCTGGATGC 698
 Db 601 GAATGAGTCGGGCTTGGAGCTGCTCAAGCTATTATCAACAGCTGTCTTCTGGATGC 660
 QY 699 GCGCTTCTCAACCCCTGCGAGGAGCTTCTTCTTCACTGGTACGACTCGCTTAC 758
 Db 661 GCGCTTCTCAACCCCTGCGAGGAGCTTCTTCTTCACTGGTACGACTCGCTTAC 720
 QY 759 TGGGGTCCCGCCAGCAGCGTGCCTTGGCTTCGAGAAGGGCAGCGTCTCTTCAACAT 818
 Db 721 TGGGGTCCCGCCAGCAGCGTGCCTTGGCTTCGAGAAGGGCAGCGTCTCTTCAACAT 780
 QY 819 CGGTGCGCTTCACAGCAGATTGGGGCGCGCAGGACCGCTCTCTGCAACCGAGGGTGC 878
 Db 781 CGGTGCGCTTCACAGCAGATTGGGGCGCGCAGGACCGCTCTCTGCAACCGAGGGTGC 840
 QY 879 CCGGCTTATGAGGCGCTTCCAGAGGGCGCTGGGGCTTCAGGCTCTGAGGAGGAACTT 938
 Db 841 CCGGCTTATGAGGCGCTTCCAGAGGGCGCTGGGGCTTCAGGCTCTGAGGAGGAACTT 900
 QY 939 CTCCATGTCGCGCCAGCCAGACATGAGCGCTGCGTCCCTCTGCGCACTGGAGCAGCTCAT 998

Db	901	CTCCATGCGCGAGCCAGACATGACGCTGGTCCCTCTCTGGCACTGGAGCAGCTCAT	960	Db	1981	TTGCAAGACCCCGCATCATCGTGGGCGACGTCCTCCGCCCCCTCTCAACTGGAGCGGAAA	2040
Qy	999	GATGGCCAGGCGCCAGGAATGTGTGTTTGGAGGCTCTTCAACCACTTGCTTCCATGGCCCGC	1058	Qy	2079	GGCCAGCAGGCGCAAGACTGGAGGCTGCCCCAGCCCTGTGCCAGTGAAGCCAGCTCC	2138
Db	961	GATGGCCAGGCGCCAGGAATGTGTGTTTGGAGGCTCTTCAACCACTTGCTTCCATGGCCCGC	1020	Db	2041	GGCCAGCAGGCGCAAGACTGGAGGCTGCCCCAGCCCTGTGCCAGTGAAGCCAGCTCC	2100
Qy	1059	CAAAGACTGCTGGCCAGCTGGCTGGCGAGAGGCGCCAGGTGGCAGCCAGTA	1118	Qy	2139	GGCCTCATCTTTGAAGCACCAGGCTGGCGC	2169
Db	1021	CAAAGACTGCTGGCCAGCTGGCTGGCGAGAGGCGCCAGGTGGCAGCCAGTA	1080	Db	2101	GGCCTCATCTTTGAAGCACCAGGCTGGCGC	2131
Qy	1119	CAGGCTAGTGACCGGACCATGCGCCAGCCACCCCTCCAGCACTAGCTGCTCTCTCTG	1178	RESULT 3			
Db	1081	CAGGCTAGTGACCGGACCATGCGCCAGCCACCCCTCCAGCACTAGCTGCTCTCTCTG	1140	ACC46698			
Qy	1179	GACTGCGCTGGTGCATGTCAAGGCGGAGTACTTCCGCTCCCTGGCCCACTTACCACTAGC	1238	ID	ACC46698	standard; cDNA; 3213 BP.	
Db	1141	GACTGCGCTGGTGCATGTCAAGGCGGAGTACTTCCGCTCCCTGGCCCACTTACCACTAGC	1200	XX	ACC46698;		
Qy	1239	CATGGCCCTTGCGACGGCTCCCGAGCGAGGAGAGCTCCCGACGACGAGCAGGT	1298	AC	ACC46698;		
Db	1201	CATGGCCCTTGCGACGGCTCCCGAGCGAGGAGAGCTCCCGACGACGAGCAGGT	1260	XX	02-JUN-2003	(first entry)	
Qy	1299	CTTCTGCGCCCGCCACCTCTCTTAAGCCCGAGGCGCTGTGTGCGCGAGAGCTGGA	1358	XX	Human	dithp biochemical pathway protein-encoding cDNA.	
Db	1261	CTTCTGCGCCCGCCACCTCTCTTAAGCCCGAGGCGCTGTGTGCGCGAGAGCTGGA	1320	XX	Human;	dithp; diagnostic and therapeutic polynucleotide; diagnosis;	
Qy	1359	GGAGCGAGGAGCTGGCGAAGGACACCTGAAGCGTGCCATCTGCTGGGGCAGGAGGCG	1418	XX	Human;	dithp; diagnostic and therapeutic polynucleotide; diagnosis;	
Db	1321	GGAGCGAGGAGCTGGCGAAGGACACCTGAAGCGTGCCATCTGCTGGGGCAGGAGGCG	1380	XX	cancer; cell proliferative disorder; autoimmune disorder;		
Qy	1419	GCTGGCGCTGACGCGCTGCTGCGCGCTGCTGCGGAGTGGA	1478	XX	inflammatory disorder; infection; hormonal disorder; metabolic disorder;		
Db	1381	GCTGGCGCTGACGCGCTGCTGCGCGCTGCTGCGGAGTGGA	1440	XX	neurological disorder; gastrointestinal disorder; transport disorder;		
Qy	1479	CTCCAGACGCTGACGCGCTGCTGCGCGCTGCTGCGGAGTGGA	1538	XX	connective tissue disorder; drug screening; proteome analysis;		
Db	1441	CTCCAGACGCTGACGCGCTGCTGCGCGCTGCTGCGGAGTGGA	1500	XX	gene therapy; antisense therapy; genotyping; transgenic animal; knock in;		
Qy	1539	CTGTGAGGCTGCGAGGCGCCCGGACATCCAGCTTAAGCCACAGAGCCAGAGGCGAG	1598	XX	disease model; toxicological testing; transcript imaging;		
Db	1501	CTGTGAGGCTGCGAGGCGCCCGGACATCCAGCTTAAGCCACAGAGCCAGAGGCGAG	1560	XX	biochemical pathway; gene; ss.		
Qy	1599	GATCCAGGCTGCTCCAGGAGAGGCGCTGACATCTTCCATCGCTGGGCGCCCTGTC	1658	XX	Homo sapiens.		
Db	1561	GATCCAGGCTGCTCCAGGAGAGGCGCTGACATCTTCCATCGCTGGGCGCCCTGTC	1620	XX	WO200297031-A2.		
Qy	1659	TGCTTCTCAGCCAGAACCGCTGGCGCTGGTGGGCGCCCTCCACCTGACCCGAGGAGA	1718	XX	05-DEC-2002.		
Db	1621	TGCTTCTCAGCCAGAACCGCTGGCGCTGGTGGGCGCCCTCCACCTGACCCGAGGAGA	1680	XX	27-MAR-2002; 2002WO-US0100056.		
Qy	1719	GGCGCGCTTGGCTCAGCTTGGGAGAGATCGCGCTGCTCCTGCTGCTGCTGCTGCTGCT	1778	XX	28-MAR-2001; 2001US-0279619P.		
Db	1681	GGCGCGCTTGGCTCAGCTTGGGAGAGATCGCGCTGCTCCTGCTGCTGCTGCTGCTGCT	1740	XX	29-MAR-2001; 2001US-0280067P.		
Qy	1779	AGGAGCCAGGCGCGCGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1838	XX	29-MAR-2001; 2001US-0280068P.		
Db	1741	AGGAGCCAGGCGCGCGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1800	XX	16-MAY-2001; 2001US-0291280P.		
Qy	1839	GCAGCCATGAGGTGGTGGAGACAGCGGAGGTGGTGGAGGTGGA	1898	XX	17-MAY-2001; 2001US-0291829P.		
Db	1801	GCAGCCATGAGGTGGTGGAGACAGCGGAGGTGGTGGAGGTGGA	1860	XX	19-JUN-2001; 2001US-0291849P.		
Qy	1899	GGCGGCGCGCGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1958	XX	20-JUN-2001; 2001US-0299776P.		
Db	1861	GGCGGCGCGCGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1920	XX	20-JUN-2001; 2001US-0300001P.		
Qy	1959	GGAGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2018	XX	(INCY-) INCYTE GENOMICS INC.		
Db	1921	GGAGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1980	XX	Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;		
Qy	2019	TTGCAAGACCCCGGATCCACGTGGGCGCAGTCCCGGCGCCCTCTCAACTGGAGCGGAAA	2078	XX	Dufour GE, Hillman JL, Yu JY, Tuason O, Yap PE, Ameshey SR;		

Claim 2; SEQ ID NO 619; 591pp; English.

The invention relates to novel human diagnostic and therapeutic polynucleotides designated dithp (ACC46080-ACC46749) and to their encoded proteins (DITHP; ABR41136-ABR41812). The invention also relates to polynucleotide sequences at least 90% identical to the dithp cDNA sequences of the invention; recombinant vectors, host cells and transgenic organisms comprising a dithp nucleic acid sequence; the recombinant production of dithp proteins; antibodies specific for dithp proteins; microarrays comprising dithp nucleic acid sequences; methods of

CC AAC78457 and AAB44240 represent sequences used in the exemplification of
CC the present invention
XX
SQ Sequence 1013 BP; 197 A; 328 C; 337 G; 151 T; 0 U; 0 Other;

Query Match 36.0%; Score 779.8; DB 3; Length 1013;
Best Local Similarity 99.6%; Pred. No. 5.7e-139;
Matches 792; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1375 GGCAGGACACCTGAAGCTGCCATCTCGGGGAGAGAGCGCTGCGGCTGCAGGCC 1434
Db |||||
QY 1435 CTGTGCGGCTCTCGCGAGGTGGACCTCTTCGGGCTGTATCTCCAGAGCTGCAG 1494
Db |||||
QY 1495 CGTCACTGGCCCAAGTATGCGAGCTCGACCTGTGAGGATGACTTCTGTAGGCTGCCGAG 1554
Db |||||
QY 1555 GCGCCGGACATCAGCTTAAGACCCAGAGCCAGAGCCAGGATGCCACGCTGTCC 1614
Db |||||
QY 1615 CAGGGGAAGGGGCTGACATCTTCCATCGGCTGGGCGCCCTGTCTGTCTCAGCCAG 1674
Db |||||
QY 1675 AACCGTGGCGGTGTGGGGCCCGTCCACCTGACCGAGGAGAGCGGCTTTGGCCTC 1734
Db |||||
QY 1735 AGCTTCGGGGAGACTGCGTCTCTATCGCTGCCGCTATCCAGGAGCCAGCGCGG 1794
Db |||||
QY 1795 GCGGCTGGGCTGAAGGAGGCGACTACATTTGTCTAGTGAATGGGAGCCATGCAGTGG 1854
Db |||||
QY 1855 TGGAGACACGCGAGGTGTGACGAGCTGAAGGCTGCGGAGAGSCGCGGCCAGCCG 1914
Db |||||
QY 1915 CAGTGGTGTCTGCTGCTGCCAGCTTAGACTGCCAGCTTGGGGACCGCGGCCGCTC 1974
Db |||||
QY 1975 CTGCTGGGCCCCAGGGGCTTCAAGGAGCCAGAGGAGCATGTTGCAGACCCCGGA 2034
Db |||||
QY 2035 TCCAGTGGGCGAGTCCCGGCGCCCTCTCACTGGAGCCGAAAGGCCAGAGGCGCAAG 2094
Db |||||
QY 2095 ACTGGAGGCTGCCCGCCGCTGTGCCCACTGAAGCCAGTCCGCTCTATCTTTGAAG 2154
Db |||||
QY 2155 CACCCAGGGTGGCGG 2169
Db |||||
QY 2169 CACCCAGGGTGGCGG 820
Db |||||

RESULT 6
AAS15828
ID AAS15828 standard; cDNA; 2100 BP.
XX
AC AAS15828;
XX
DT 16-JAN-2002 (first entry)

XX
DE
XX
KW Human; prostate specific protein; PSL22; prostate cancer;
KW benign prostatic hyperplasia; acute prostatitis; testicular cancer;
KW cryptorchidism; testicular disorder; proliferative disorder; lymphoma;
KW leukaemia; melanoma; ovarian cancer; breast cancer; pancreatic cancer;
KW liver cancer; lung cancer; cytostatic; ss.
XX
OS Homo sapiens.
XX
XX
FH Key Location/Qualifiers
FT CDS 1..2061
FT /*tag= a
FT /product= "PSL22"
XX
XX WO200172962-A2.
XX
XX 04-OCT-2001.
XX
XX 23-MAR-2001; 2001WO-US009410.
XX
XX 24-MAR-2000; 2000US-0191929P.
XX
XX (SAAT/) SAATCIOGLU F.
XX
XX Saatcioglu F;
XX
XX WPI: 2001-662926/76.
XX P-PSDB; AAU10192.
XX
XX New polynucleotide for the diagnosis, prevention and treatment for
XX prostate and testis disorders, particularly prostate cancer, comprises
XX prostate-specific or testis-specific nucleic acids.
XX
XX Claim 4; Fig 14B; 114pp; English.
XX
XX The invention relates to substantially pure prostate-specific or testis-
XX specific polypeptides and the nucleic acids encoding them. Also included
XX are vectors and host cells expressing the proteins, a transgenic animal
XX expressing the protein, antibodies against the proteins, probes for
XX detecting the nucleic acids, antisense molecules for the nucleic acids
XX and methods of isolating modulators of the proteins. Compounds that
XX modulate the prostate specific or testis specific polypeptide are useful
XX to diagnose, prevent or treat disorders of the testis or prostate
XX particularly prostate cancer, benign prostatic hyperplasia, acute
XX prostatitis, testicular cancer, cryptorchidism, undescended, retractile,
XX ascending or vanished testis. Other proliferative disorders for which the
XX modulators may be used include lymphoma, leukaemia, melanoma, ovarian
XX cancer, breast cancer, pancreatic cancer, liver cancer and lung cancer.
XX The present sequence encodes a prostate specific protein, PSL22

SQ Sequence 2100 BP; 516 A; 569 C; 584 G; 431 T; 0 U; 0 Other;

Query Match 22.9%; Score 496.4; DB 5; Length 2100;
Best Local Similarity 56.8%; Pred. No. 4.7e-85;
Matches 977; Conservative 0; Mismatches 731; Indels 12; Gaps 3;

QY 219 GGGCTGTGACTCCCTGACGCGAGATCCAGTGGCGGCGAGCTGCAGAGCCGAGGCCAGAT 278
Db |||||
QY 69 GGGCTGTAAATCCCTTGCAAAACCGGCGGAGTAAATTCAGAAATCAAGAGCTGCTT 128
Db |||||
QY 279 TCACGAGCAGATTGACAAAGGAGCTGCAGATGCGGACGGGCTGAGACCTCTACAGAGC 338
Db |||||
QY 129 GAATCAGCAGATCTGAAAGCCGTGCGGATGAGGACCGGAGAAACCTTCTGAAAGT 188
Db |||||
QY 339 CACCAACAACACCGGCTGAGAGAGAGCGTCCGCTGAGCTGAGCTACGCTCACTCAA 398
Db |||||
QY 189 GGCCACAAACTCAAAGGTGCGGAGCAAGTGGCGCTGGAGCTGAGCTTGTCAACTCAGA 248
Db |||||
QY 399 CCGCAGCTGCTGAAGGAGGAGCTGGAGGAGCTCAGCGGTGGCGTGGACCCCTGCCCGCA 458
Db |||||
QY 249 CCGCAGATGCTCAAGGAAGAGCTGGAGGAGCTGAACATCTCGGTGGCGCTCTATCAGAA 308
Db |||||

QY	459	TGGAGCGAAGCTGTCACTGTGTCCTCCCATGATCCCTTGGGCTTGAAGAGACCAAGAGCT	518	QY	1533	TGACTTCTGTGAGGCTGCCGAGGCCCGGACATCCAGCTTAAGACCCACGAGCCAGA	1592
Db	309	CAAGAGGAGGCAATTTACGATTCCTCCCTGATTCCTCTTGGCTTGAAGAAACGAAAGACGT	368	Db	1383	TGACCTGTCTGAACCTGATCGACGCCGCCAGTGTGTGCTTAAACTGAGCAAGAGTTGA	1442
QY	519	GGACTGGTCTACACCGCTGAAGAGCTGATCTCAGTGCACCTTTGGAGAGACGGCGCTC	578	QY	1593	GGCCAGGATGCCACGGCTGTCCAGGGGAAGGGCTTGACATCTTCCATCGGCTGGGGCC	1652
Db	369	CGACTTTTGAGTGGTCTCAAGGATTTTATCTTGGAACTTACAGTGAAGATGGCTATT	428	Db	1443	CATTATATTGCCGCCAGTTCTCCAAGCTGACAGTCAAGGACTTCTCCAGAGCTGGCCC	1502
QY	579	CTACGAGGAGAAATCAGGAGCTGGAGGCCCTGCGGACAGGCCATCGGACCCCGAGCCG	638	QY	1653	CGTGTCTGTGTTCTCAGCCAAGAACCGGTGGGGCTGGTGGGGCCCGTCCACCTGACCCG	1712
Db	429	ATATGAGATGAATTCAGATCTTATGATCTGAGACAGCTTGTGCGACGCTAGCCG	488	Db	1503	CTTATCTGTGTTTTTCGCTTAAACAGCGGTGGAGCCCTCCCTCGAAGCATCGCTTCACTGC	1562
QY	639	GAATGAGTGGGCTGTGAGTGTCTCAGAGCTTATTAACACAGCTGTGCTTCTGATGC	698	QY	1713	AGGAGAGGGCGCTTTGGCTTCAAGCTTCCAGGAGACTCGCTGTCTCTATCCTGCTCCGT	1772
Db	489	GGATGAGGCGGGGTGAACTGTCTGATGACATATCTATCCAGCTGGGCTTTGTGAGAG	548	Db	1563	AGAAGAAGGGACTTGGGGTTACCTTTGAGAGGAAACGCCCGCTTCAGGTTCACTTCCT	1622
QY	699	CGCTTCTCTACCCCTGCCAGGAGCCTCGGGCTCTTCTCCACTGGTACGACTCGCTTAC	758	QY	1773	CATTCCAGGAGCACCGCCGCGGCTGGCTGAAAGGAGGGCGACTACATTGTGTCACT	1832
Db	549	TCGATTTCTTCCCGCCACACGGAGATGGGACTCTCTGTTCACTGGTATGACTCTCTAC	608	Db	1623	GGATCCTTACTGCTCTGCTCGGTGGAGAGCCCGGGAAGAGATTTATTTGTCTCCAT	1682
QY	759	TGGGCTTCCCGCCACAGAGCGTGCCTTGGCTTTGAGAGGGGAGCGCTTCTTCAACAT	818	QY	1833	GAATGGCGACCATGACAGTGTGTGGAGACACGGCGAGGTGTGACGAGCTGAAGGCTGC	1892
Db	609	TGGGTTCCCGTCAAGCCAGCAGAACCTGCTGTGGAGAGGCCAGTCTCTGTTCAACAC	668	Db	1683	TCAGCTGTGGATTGAAGTGGCTGACGCTGAGTGAGGTTATGAAGCTGCTGAAGAGCTT	1742
QY	819	CGTGTCCCTCCACAGCAGATTTGGGCGCGCCAGGACCGTCTCTGACCCAGGCTGCCG	878	QY	1893	GGGAGAGGGCGGCGCCAGCCTGACAGTGTGTGCTGCTGCTG 1932	
Db	669	TGGGGCCCTCTACACCAGATTTGGAGCCCGTGCATCGGACAGCGCTGGGCTGGGA	728	Db	1743	TGGGAGGACGAGATCGAGATGAAGTGAAGTGTGTGAGCCTCCTG 1782	
QY	879	CGCGCTATGGAGCCTTCCAGAGGGCGCTGGGGCTTCAAGCTCTCTGAGGAGAACTT	938	RESULT 7			
Db	729	GAGTGCATAGATGCCCTTTCAGAGCGCCGAGGGGTTTAAATTTACCTGAAAGACATTT	788	AA515827			
QY	939	CTCCATCTCCAGTTCAGACATGAGCGTGTGCTTCCCTTCCGCACTGGAGAGCTCAT	998	ID	AA515827	standard; cdna; 3526 BP.	
Db	789	TACCATCTCCAGTTACGACATGAGCCCTGCATGCTCAGCGTGTCTGCTCAAAATGAT	848	XX	AA515827;		
QY	999	GATGGCCAGGCCAGGAATGTGTGAGGGCTCTCACCACCTGCTCCATGGGCC	1058	XX	16-JAN-2002	(first entry)	
Db	849	GCTTGCACAAAGGAGCGGTTTGAAGAAATCAGCTTCTCTG-----GATCCG	902	XX	Human ORF for prostate specific protein PSL22.		
QY	1059	CCAGACTGCTGCGCCAGCTGCGCTGGCGCAGGAGGCCCGCCAGGTGGAGCGAGTA	1118	XX	Human; prostate specific protein; PSL22; prostate cancer;		
Db	903	GAATGAATCTTCATGCTGTGAGGTGCTCAGAGGCTCTAAGGTGGAGAGTCTA	962	XX	benign prostatic hyperplasia; acute prostatitis; testicular cancer;		
QY	1119	CAGGCTAGTGCAGCGACATGGCCACCCGCTCCAGCTACGCTACGTCGCTCTCTG	1178	XX	cryptorchidism; testicular disorder; proliferative disorder; lymphoma;		
Db	963	CCACAGCTACAGCAGCCATGAGCCAGCGCCGCTGAAAGAGAAACATCCCTTACTCTCTG	1022	XX	leukemia; melanoma; ovarian cancer; breast cancer; pancreatic cancer;		
QY	1179	GACTGCTGTGTGATGTCAAGCGGAGTACTTCCGCTCCCTGGCCCACTACCGTAGC	1238	OS	Homo sapiens.		
Db	1023	GGCCAGCTTAGCTGCGTGAAGGCCCACTACAGCGGCGCTGACACCTTGGCCCACTTCACTGC	1082	XX	Key	Location/Qualifiers	
QY	1239	CATGCGCTCTGCGACGCTCCCGACGACGAGGAGGCTCCCGACGACGAGCAGT	1298	XX	1. .2061		
Db	1083	CATCTCTCATGACACAGGTGAAGCCAGGACGGATCTGGACCAACGAGGAGAGTG	1142	FT	/tag= a		
QY	1299	CTTCTGACGCG-----CCCCACCTCTTAAGCCCCGAGGCCCTGTGC--TGCCCGCAGA	1352	FT	/product= "PSL22"		
Db	1143	CCTGTCCAGCTCTACGACCAATGCCAGAGGGGTGACACCTTGGCCCACTGAAGAA	1202	XX	WO200172962-A2.		
QY	1353	GCTGAGAGCGCAGGAGCTTGGCAAGCAGCAGCTGAAGCGTGCCTCTGGGCGAGA	1412	XX	04-OCT-2001.		
Db	1203	TGATCAGAGCGCCGACAGCTGGGAAGTCCCACTTGGCAGAGCATGGCTCATCAGA	1262	XX	23-MAR-2001; 2001WO-US009410.		
QY	1413	GGAGCGCTGCGGCTGCACGCTGTGCGCGCTCTGCGCGAGGTGACCTGCTCGGGC	1472	XX	24-MAR-2000; 2000US-0191929P.		
Db	1263	GGAGTCTGGTGGGAGCGAGCTCTGCAAGAGCTGCGAGCAITTAGAGTGTACAGAA	1322	XX	(SAAT/) SAATCIOGLU F.		
QY	1473	TGTGATCTCCAGAGCTGACGCTCACTGCGCAAGTATGCGGAGCTGACCGGTGAGA	1532	XX	Saatcioglu F;		
Db	1323	GCTGTGTGTGCGCAGCAGGAAGCTCCCGGCTCAGTACGCCCGCAGCAGGAGGAGA	1382	XX	WPI; 2001-662926/76.		
				XX	P-PSDB; AAU10192.		
				XX	New polynucleotide for the diagnosis, prevention and treatment for		
				XX	prostate and testis disorders, particularly prostate cancer, comprises		
				XX	prostate-specific or testis-specific nucleic acids.		
				XX	Claim 5; Fig 14A; 114pp; English.		

The invention relates to substantially pure prostate-specific or testis-specific polypeptides and the nucleic acids encoding them. Also included are vectors and host cells expressing the proteins, a transgenic animal expressing the protein, antibodies against the proteins, probes for detecting the nucleic acids, antisense molecules for the nucleic acids and methods of isolating modulators of the proteins. Compounds that modulate the prostate specific or testis specific polypeptide are useful to diagnose, prevent or treat disorders of the testis or prostate particularly prostate cancer, benign prostatic hyperplasia, acute prostatitis, testicular cancer, cryptorchidism, undescended, retractile, ascending or vanished testis. Other proliferative disorders for which the modulators may be used include lymphoma, leukaemia, melanoma, ovarian cancer, breast cancer, pancreatic cancer, liver cancer and lung cancer. The present sequence encodes a prostate specific protein, PS122

Sequence 3526 BP; 992 A; 803 C; 836 G; 895 T; 0 U; 0 Other;

Query Match 22.9%; Score 496.4; DB 5; Length 3526;
Best Local Similarity 56.8%; Pred. No. 5e-85;
Matches 977; Conservative 0; Mismatches 731; Indels 12; Gaps 3;

```

QY 219 GGGCTGTGACTCCCTGACGCGAGATCCAGTGGCGGCCAGCTGCAGAGCCGAGGCGCCAGAT 278
   |||||
Db 69 GGGCTGTAAATCCCTTGCACAAACCGCGCGAGTAAATTGCAGATCAAGAGCTGCTT 128

QY 279 TCACCAGCAGATTGACAGAGAGCTGCAGATGCGGAGCGGCGCTGAGAACCTTCAAGAGC 338
   |||||
Db 129 GAATCAGCAGATCCCTGAAGCCGTCGGATGAGGACCGGAGCGGAAACCTTCTGAAAGT 188

QY 339 CACCAGCAACACCGGGTGCAGAGACGCTGCCCTGGAGCTGAGCTAGCTCAACTCCAA 398
   |||||
Db 189 GGCACAAACTCAAGGTGCGGAGCAAGTGGGCTGAGAGTACCTTCGTCAACTCAGA 248

QY 399 CTTGAGCTGCTGAAGGAGAGCTGGAGAGCTCAGCGGTGGCGTGGACCCCTGGCGGCA 458
   |||||
Db 249 CTTGAGAGTGTCAAGGAGAGCTGGAGGGCTGAACATCTCGGTGGCGTCTATCAGAA 308

QY 459 TGGGAGCGAAGCTGTCACTGTCCCATGATCCCTGGCGCTGAGAGAGCAAGAGAGCT 518
   |||||
Db 309 CACAGAGAGGCAATTACGATTCCTCTGCGCTGAGGAGGAAAGAGAGCT 368

QY 519 GGACTGTGTACACCGCTGGAAGGAGCTGATCTCAGTGCACATTTGGAGAGGACGCGCCTC 578
   |||||
Db 369 CGACTTTGACGTGCTCCTCAGGATTTTATCTGGAACTTACAGTGAAGATGGCTATT 428

QY 579 CTAGAGCAGAAATCAGGAGAGCTGGAGCCCTGCGCAGGCAATGCGGACCCCGAGCCG 638
   |||||
Db 429 ATATGAAGATGAAATTGCAGATCTTATGGATCTGAGACAAAGCTTTCGACGCTAGCCG 488

QY 639 GAATGAGTCGGGCTGGAGCTGCTACAGCCTATTACAAACAGCTGTCTTCCTGGATGC 698
   |||||
Db 489 GGATGAGGCGGGTGGAACTGCTGATGACATCTTATCCAGCTGGGCTTTGTGAGAG 548

QY 699 GCGCTTCTCAACCCCTGCGAGGAGCTCGGGCTCTTCTTCCACTGGTACGACTCGCTTAC 758
   |||||
Db 549 TCGATTTCTTCCGCGCCACACGCGAGATGGGACTCTGTTCACCTTGTATGACTCTCTCAC 608

QY 759 TGGGGTCCCGGCCAGCAGCTGCGCTTTCGAGAAGGCGAGCGTCTCTTCAACAT 818
   |||||
Db 609 TGGGGTTCCGGTCAAGCAGCAGAACTGCTGCTGAGAAGGCCAGTGTCTCTGTTCAACAC 668

QY 819 CGGTGCCCTCCACAGCAGATGGGCGCGCAGAGCCGCTCTGCAACGAGGAGTGGCCG 878
   |||||
Db 669 TGGGGCTCTTACACCCAGATGGGACCCGCTGCGATCGGAGACGAGGCTGGGCTGGA 728

QY 879 CCGGCTATGAGGCGCTTCCAGAGGCGCGTGGGCGCTTACGCTCTGAGGAGAACTT 938
   |||||
Db 729 GAGTGCCATAGATGCTTTTCAGAGAGCGCAGGGGTTTAAATACCTTGAAGACACATT 788

QY 939 CTCCTCATGCGCGAGCCAGACATGAGCGCTGCGTCCCTCTGCGCACTGGAGCAGCTCAT 998
   |||||
Db 789 TACCCATACTCCAAGTTACGACATGAGCCCTGCCATGCTCAGCGTCTCGTCAAAATGAT 848

```

```

QY 999 GATGGCCCAAGGCCCAAGGAATGTGTGTTTGAAGGCGCTTCTCACCACCTGCTCCATGCCCCC 1058
   |||||
Db 849 GCTTGCAAGCCCAAGAAAGCGTGTGTTGAGAAATCAGCCTTCTCTG-----GGATCCG 902

QY 1059 CCAAGACTGCTGGGCCAGCTGCGCTGGCGAGAGGCGCCGAGGTGGAGCCGAGTA 1118
   |||||
Db 903 GAATGAATTTCTTCATGTGTTGAAGTGGCTCAGGAGGCTGTAAAGTGGAGAGGTCTA 962

QY 1119 CAGGCTAGTGCACCGGACCATGGCCCGAGCCACCGCTCCACAGACTAGTGCCTGCTCTCTG 1178
   |||||
Db 963 CCAACAGCTTACAGCAGCATGAGCAGCGCGCGGTGAAGAGAAATCCCTACTCTCTG 1022

QY 1179 GACTGCGCTGGTGCATGTCAAGCGCGAGTACTTCCGCTTCCCTGGGCCACTACCACTGATG 1238
   |||||
Db 1023 GGCAGCTTAGCTGCTGTAAGGCCCAACCACTACGCGGCGCTGGCCCACTACTTCTCACTGC 1082

QY 1239 CATGGCCCTCTGCGAGGCTCCCGAGCGACCGAGGAGAGCTCCCGACGACGAGAGGT 1298
   |||||
Db 1083 CATCTCTCTCATCGACCAACCAAGGTGAAGCCAGGACGAGCTCTGGACCAACGAGGAAAGTG 1142

QY 1299 CTTCTGCGAGC-----CCGCCACTCTCTTAAGCCCGGAGGCCCTGTGC--TGCCGACGA 1352
   |||||
Db 1143 CTTGCTCCAGCTCTACGACCAATGCCAGAGGGGTGACACCTTTGGCCCACTGAAAGAA 1202

QY 1353 GCTGAGGAGCGCAGGAGCTTGGCAAGGCACACCTGAAGCGTGCCTCTGGGCGACGA 1412
   |||||
Db 1203 TGATCAGCAGCGCCGACAGCTGGGGAAAGTCCACACTTTCGAGAGCCATGGCTCATCAGA 1262

QY 1413 GGAGCGCTGCGGCTGACAGCGCTGTGCGGCTCTTCGCGAGGAGTGCCTGCTTCGGGCG 1472
   |||||
Db 1263 GGAGTCGCTGGGAGGCGAGCTCTTCAAGAGCTGCGGAGCATTTGAGGTGCTACAGAA 1322

QY 1473 TGTGATCTCCAGACGCTGACAGCGCTCACTGCGCAAGTATGCGAGCTCGACCGTGAGGA 1532
   |||||
Db 1323 GGTGCTGTGTCGCGACAGGAAAGCTTCCGGCTCAGGTACGCCCGACGACGAGGAGAGA 1382

QY 1533 TGACTTCTGTGAGGCTGCCAGGCGCCCGGAATCAAGCTTAAGACCCACGAGAACCCAGA 1592
   |||||
Db 1383 TGACTGTGTAACCTGATCGACGCGCCCGCAGTGTGTTGCTTAAACTGAGCAAGAGTTGA 1442

QY 1593 GGCCAGAGTGCACGCTGTCAGAGGAGGCGGCTGAGTCTTCCATTCGGCTGGGCGC 1652
   |||||
Db 1443 CATATATATGCGCCAGTTCTCCAAGCTGACAGTCAAGGAGTCTTCTCCAGAAAGCTGGGCGC 1502

QY 1653 CCTGTCTGTGTTCTCAGCAAGAACCGGTGGCGGCTGGTGGGCGCGCTCCACCTGACCCG 1712
   |||||
Db 1503 CTTATCTGTGTTTTCGGCTAACAGCGGTGACGCGCTCTCGAAGCATCCGCTTCACTGCG 1562

QY 1713 AGGAGAGGCGGCTTTCGCTCAAGCTTTCGGGAGACTGCGCTGTCTCATCGCTGCGCGT 1772
   |||||
Db 1563 AGAAGAAGGGGACTTGGGGTTTCACTTGAGAGGGAACGCCCCCGTTCAAGTTCACTTCT 1622

QY 1773 CATTCAGGAGGAGCCAGCGCGCGGCTGGCTCAAGAGGAGGCGGCTACATTTGTGCTAGT 1832
   |||||
Db 1623 GGATCTCTTATGCTCTGCTGCTGGTGGAGGAGCCCGGAGAGGAGATATATTTGTCTCAT 1682

QY 1833 GAATGGGCGAGCCATGCAAGTGGTGGAGACAGCGCGAGGTGCTGACGAGGCTGAAGGCTGC 1892
   |||||
Db 1683 TCAGTTGTGGATGTAAGTGGCTGACGCTGAGTGAAGTTATGAAGCTGCTGTAAGAGCTT 1742

QY 1893 GGGAGAGCGGCGCCAGCTGCAAGGTGGTGTGCTGCTG 1932
   |||||
Db 1743 TGGCGAGGACGAGATCGAGATGAAGTGTGCTGAGCCTCTG 1782

```

RESULT 8

ABX97183

ID ABX97183 standard; cDNA; 2109 BP.

XX ABX97183;

XX 20-MAY-2003 (first entry)

XX

DE Human NOV126b cDNA.
XX NOVX; cytostatic; cardiant; antiarteriosclerotic; antiasthmatic; cancer;
KW hypotensive; cardiomyopathy; bronchial asthma; gene therapy; vaccine;
KW human; gene; ss.
XX Homo sapiens.
XX WO200272757-A2.
XX 19-SEP-2002.
PF 08-MAR-2002; 2002WO-US006908.
XX 08-MAR-2001; 2001US-0274101P.
PR 08-MAR-2001; 2001US-0274194P.
PR 08-MAR-2001; 2001US-0274281P.
PR 08-MAR-2001; 2001US-0274322P.
PR 09-MAR-2001; 2001US-0274849P.
PR 12-MAR-2001; 2001US-0275235P.
PR 13-MAR-2001; 2001US-0275578P.
PR 13-MAR-2001; 2001US-0275579P.
PR 13-MAR-2001; 2001US-0275601P.
PR 14-MAR-2001; 2001US-0276000P.
PR 16-MAR-2001; 2001US-027676P.
PR 19-MAR-2001; 2001US-0276994P.
PR 20-MAR-2001; 2001US-0277239P.
PR 20-MAR-2001; 2001US-0277321P.
PR 20-MAR-2001; 2001US-0277327P.
PR 21-MAR-2001; 2001US-0277791P.
PR 22-MAR-2001; 2001US-0277833P.
PR 23-MAR-2001; 2001US-0278152P.
PR 26-MAR-2001; 2001US-0278894P.
PR 27-MAR-2001; 2001US-0278999P.
PR 27-MAR-2001; 2001US-0279036P.
PR 28-MAR-2001; 2001US-0279344P.
PR 30-MAR-2001; 2001US-0277338P.
PR 30-MAR-2001; 2001US-0279995P.
PR 30-MAR-2001; 2001US-0280233P.
PR 02-APR-2001; 2001US-0280802P.
PR 02-APR-2001; 2001US-0280822P.
PR 02-APR-2001; 2001US-0280900P.
PR 04-APR-2001; 2001US-0281194P.
PR 13-APR-2001; 2001US-0283675P.
PR 30-APR-2001; 2001US-0287424P.
PR 02-MAY-2001; 2001US-0288066P.
PR 03-MAY-2001; 2001US-0288342P.
PR 03-MAY-2001; 2001US-0288528P.
PR 15-MAY-2001; 2001US-0291190P.
PR 16-MAY-2001; 2001US-0291099P.
PR 16-MAY-2001; 2001US-0291240P.
PR 30-MAY-2001; 2001US-0294485P.
PR 31-MAY-2001; 2001US-0294889P.
PR 31-MAY-2001; 2001US-0299027P.
PR 18-JUN-2001; 2001US-0299303P.
PR 19-JUN-2001; 2001US-0299310P.
PR 10-JUL-2001; 2001US-0304354P.
PR 31-JUL-2001; 2001US-0309198P.
PR 16-AUG-2001; 2001US-0312903P.
PR 10-SEP-2001; 2001US-0318462P.
PR 12-SEP-2001; 2001US-0318770P.
PR 27-SEP-2001; 2001US-0325430P.
PR 27-SEP-2001; 2001US-0325681P.
PR 18-OCT-2001; 2001US-0330380P.
PR 31-OCT-2001; 2001US-0335301P.
PR 14-NOV-2001; 2001US-0332172P.
PR 14-NOV-2001; 2001US-0332271P.
PR 14-NOV-2001; 2001US-0332272P.
PR 14-NOV-2001; 2001US-0333184P.
PR 14-NOV-2001; 2001US-0333272P.
PR 21-NOV-2001; 2001US-0332094P.
PR 03-DEC-2001; 2001US-0337426P.

PR 03-DEC-2001; 2001US-0338092P.
PR 04-DEC-2001; 2001US-0337185P.
PR 03-JAN-2002; 2002US-0345705P.
PR 07-MAR-2002; 2002US-00092900.
XX (CURA-) CURAGEN CORP.
XX Padigaru M, Spytek KA, Shenoy SG, Taupier RJ, Pena CEA, Li L;
PI Zerhusen BD, Gusev V, Ji W, Gorman L, Miller CE, Kekuda R;
PI Pattarajan M, Gangolli E, Vernet CAM, Guo X, Tchernev V;
PI Fernandes ER, Casman SJ, Malyankar UM, Gerlach V, Liu Y, Anderson D;
PI Spaderna SK, Catterton E, Burgess C, Leite M, Zhong H, Alsobrook JB;
PI Lepley DM, Rieger DK;
XX WPI; 2002-723332/78.
DR P-PSDB; ABU65216.
XX NOVX polypeptides and polynucleotides, useful for preventing or treating
PT a disorder associated with aberrant NOVX expression or activity e.g.,
PT cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial
PT asthma.
XX Claim 13; Page 471; 1103pp; English.
XX This invention describes novel human NOVX polypeptides which have
CC cytostatic, cardiant, antiarteriosclerotic, antiasthmatic and hypotensive
CC activity. Pharmaceutical compositions comprising the NOVX proteins or
CC nucleic acid molecules or NOVX antibodies are useful for preventing or
CC treating a disorder associated with aberrant NOVX expression or activity
CC e.g. cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial
CC asthma. The products of the invention can be used for gene therapy or in
CC a vaccine. ABX97008-ABX97185 are cDNA fragments amplified and isolated by
CC the PCR primers and probes represented in ABX13460-ABX13462 and ABX97186-
CC ABX97593. ABX97008-ABX97185 encode the NOVX proteins described in
CC ABU65041-ABU65218
XX SQ Sequence 2109 BP; 518 A; 574 C; 585 G; 432 T; 0 U; 0 Other;
Query Match 22.8%; Score 494.8; DB 6; Length 2109;
Best Local Similarity 56.7%; Pred. No. 9.4e-85;
Matches 976; Conservative 0; Mismatches 732; Indels 12; Gaps 3;
QY 219 GGGCTGTGACTCCCTGACGCGAGATCCAGTCGCGCGAGCTCGAGCGCGAGGGGCCAGAT 278
DB 79 GGGCTGTATCCCTTTCACAAACCGCGCGAGTAAATTCAGAAATCAAGAGCTGCTTT 138
QY 279 TCACGACGATTGACAGAGAGCTGCAGATCGGACGGCGCTGAGAACCTTACAGAGC 338
DB 139 GAATCAGCAGATCCTGAAAGCCGTCGGGATGAGGACCGGAGCGGAAACCTTCTGAAAGT 198
QY 339 CACCAGCAACACCGGCTGAGAGACGCTCGCCCTGGAGCTGAGCTACCTCAACTCAA 398
DB 199 GGCCACAAATCAAGTGGCGGAGCAAGTGGCGCTGAGCTGAGCTTCGTAACACTCAGA 258
QY 399 CTTGCAGCTCTGAAGGAGGAGCTGGAGGAGCTCAGCGGTGGCGCTGAGACCTTGGCCGCGCA 458
DB 259 CTTGCAGATGCTCAAGGAGAGCTGGAGGGGCTGAACATCTCGGTGGCGCTTATCAGAA 318
QY 459 TGGGAGGAGAGCTGCTACTGTCCCATGATCCCTCGGCTGAAGGAGACCAAGGAGCT 518
DB 319 CACAGAGGAGGCAATTTACGATTCCTTGGCTTGAAGAAACGAAAGAGCT 378
QY 519 GGACTGGTCTACACCGCTGAAGGAGCTGATCTCAGTGCATCTTTGGAGAGGACGCGCTC 578
DB 379 CGACTTTGCAGTCTCTCAAGGATTTATCTGGAACATTTACAGTGAAGATGGCTATTT 438
QY 579 CTACGAGGCGAATTCAGGAGCTGGAGGCGCTCGGCGAGCCATCGGACCCCGCCGCG 638
DB 439 ATATGAAGATGAATTCAGATCTTATGATCTGAGACAACTTGTGCGAGCGCTAGCCG 498
QY 639 GAATGAGTGGGCTGGAGCTGCACAGCGCTATTACAAACAGCTGTGCTTCCTGGATGC 698
DB 499 GGATGAGCGCGGCTGGAATGCTCTGATGACATATCTTATCCAGCTGGCTTTGTGAGAG 558

QY 699 GGGCTTCTCACCCCTGCCAGAGCCCTCGGCTCTTCTTCCACTGATGACACTGGCTTAC 758
Db 559 TGAATTTCTTCCCGCCCAACAGCGGAGATGGGACTCTCTGTTTCACTGGTATGACTCTCTAC 618
QY 759 TGGGGTCCCGGCCAGCAGCGTGCCTTGGCTTTCGAGAAAGGCGAGGTTCTTCTCAACAT 818
Db 619 CGGGTTCGGTTCAGCAGCAGAACCTTGTCTGGAGAAAGCCAGTGTCTCTGTTCACAC 678
QY 819 CGGTGCCCTCCACACACAGATTGGGCGCGCCAGGACCGCTCTGACCCGAGGGTCCCG 878
Db 679 TGGGGCCCTCTACACCAGATTGGGACCCGGTGGATGCGACGACGACGCGTGGTGA 738
QY 879 CGCGCTATGGAGGCTTCCAGAGGCGCGTGGGCTTTCAGCCTCTGAGGAGAACTT 938
Db 739 GAGTGCCATAGATGCTTTTCAGAGAGCCCGAGGGTTTAAATTAACCTGAAAGACACTT 798
QY 939 CTCCATGCGCGAGCCAGACATGAGCGTGTGCTCCCTCTGCGACTGCGAGCTCAT 998
Db 799 TACCCATCTCCAAGTTACGACATGAGCCCTGCGCTCAGCTGCTCGTCAAAATGAT 858
QY 999 GATGGCCAGGCCAGGAATGTGTGTGAGGCGCTTCAACCACTGCTCCATGCGCCC 1058
Db 859 GCTTGCAAGCCCAAGAAAGCGTGTTCAGAAATCAGCTTCTCTG-----GGATCG 912
QY 1059 CCAAGACTGCTGGCCAGCTGCGCTGGCTGGCAGAGGCGCCAGGTGGAGCGAGTA 1118
Db 913 GAATGAATTTCTATGCTGGTGAAGGTGGCTCAGGAGGCTGCTAAGGTGGAGAGTCTA 972
QY 1119 CAGGCTAGTGCACCGGACCATGCGCCAGCCACCCGTCACGACTAGTGCCTGTCTCTG 1178
Db 973 CCAACAGCTTACAGCGACCATGAGCAGCGCGCGGTGAAGAGAACATCCCTACTCTCTG 1032
QY 1179 GACTGCCCTGTGCGATGTCAAGCGCGAGTATTCCGCTCTCTGGCCACTACCACTGAC 1238
Db 1033 GCGCAGCTTAGCTGTGTAAGGCCCAACCACTACGCGGCTGTGGCCACTACTTCACTGC 1092
QY 1239 CATGGCCCTCTGCGAGCGTCCCCAGCGACCCAGGAGAGCTCCCCAGCAGCAGAGGT 1298
Db 1093 CATCTCTCTATGACACCAACAGGTGAAGCCAGGCACTGTGGACCAACAGGAGAAAGT 1152
QY 1299 CTCTCTGAGC-----CCCCACCTCTCTTAAGCCCGAGGCGCTGTGC--TGCCGACGA 1352
Db 1153 CTGTCCCAGCTCTAGACCATGCGCAGAGGGGTGACACCTTGGCCACACTGAAGAA 1212
QY 1353 GGTGAGAGAGCCAGCGCTTGGCAAGCACAACCTGAAGCGTGCATCTTGGGCGACGA 1412
Db 1213 TGATCAGCAGCGCCGACAGCTGGGGAAGTCCCACTTGCGCAGCCTATGGCTCATCGA 1272
QY 1413 GAGGCGCTGGGCTGCACGCCCTGTGCGCGCTCTGCGGAGGTGGACCTGCTTCGGGC 1472
Db 1273 GAGTCTGTGGGAGGCAAGCTCTGCAAGAAAGCTGCGGAGCAATTGAGGTGTACAGAA 1332
QY 1473 TGTGATCTCCAGAGCTGCGAGCGCTCACTGCGCAAGTATGGGAGCTCGACCGTGAGGA 1532
Db 1333 GTGTCTGTGTCGCGACAGGAAGCTTCCGGCTCAGCTACGCCCGCAGCACAGGAGGGA 1392
QY 1533 TGACTTCTGTGAGGTGCGGAGGCCCGGACATCCAGCTTAAGACCCACCAAGGCCAGA 1592
Db 1393 TGACCTGTGAACCTGATCGACGCCGCCCCAGTGTGTTGTTGCTAAATCTGAGCAAGAGTTGA 1452
QY 1593 GGCCAGGATGCCAGCCTTCCAGGGAAGGGGCTTACATCTTCATTCGCTGGGGCC 1652
Db 1453 CATTAATATGCCCCAGTTCTTCAAGCTGACAGTCAAGGACTTCTTCCAGAAAGCTGGGCC 1512
QY 1653 CTGTCTGTGTTCTCAGCAAGAACCGGTGCGGCTGTGGGCGCCGCTGACCTGACCCG 1712
Db 1513 CTTATCTGTGTTTCGGGTAAACAGCGGTGAGCGCTCTCTGAAGCAATCCGCTTCACTGC 1572
QY 1713 AGGAGGCGCGCTTGGCTTCCAGCTTCCGGGAGACTTCGCTGTCTCTCATCGCTGCCGT 1772
Db 1573 AGAAGAGCGGACTTGGGGTTTCACTTGTGAGAGGGAAGCCGCCCGCTTCAAGTTCACTTCT 1632

QY 1773 CATTCAGGAGCAGGCGCGGCTGCGCTGAAGAGGGCGCACTACATTGTGTCACT 1832
Db 1633 GGATCCTTACTGCTCTGCTCGCTCGTGGCAGGAGCCCGGAGAGATATATTGTCTCAT 1692
QY 1833 GAATGGGAGCCATGACAGGTGTGTGGAGACACCGGAGGTGTGACCGAGCTGAAGGCTGC 1892
Db 1693 TCAGCTTGTGGATTGTAAAGTGGCTGACGCTGAGTGAGGTATGAAGCTGCTGAAGAGCTT 1752
QY 1893 GCGAGAGGCGGGCCGAGCGCTGACAGTGTGTGCTGCTGCTG 1932
Db 1753 TGGCAGGAGCAGATCAGATGAAGTCGTGAGCCTCTCTG 1792

RESULT 9
ABS71856

ID ABS71856 standard; cDNA; 2061 BP.

XX ABS71856;

XX DT 02-DEC-2002 (first entry)

XX Human GTP-Rho binding protein 2 open reading frame.

XX Human; ss; GTP-Rho binding protein 2; GRBP2; chromosome 19q12; oncogene;
tumour; liposarcoma; ichthyosis congenita III;
benign familial infantile convulsion; gene therapy.

XX OS Homo sapiens.

XX PN EP1231216-A2.

XX PD 14-AUG-2002.

XX PF 17-JAN-2002; 2002EP-00001026.

XX PR 30-JAN-2001; 2001WO-US000663.

XX PR 30-JAN-2001; 2001WO-US000664.

XX PR 30-JAN-2001; 2001WO-US000665.

XX PR 30-JAN-2001; 2001WO-US000666.

XX PR 30-JAN-2001; 2001WO-US000667.

XX PR 30-JAN-2001; 2001WO-US000668.

XX PR 30-JAN-2001; 2001WO-US000669.

XX PR 29-JUN-2001; 2001US-00895040.

XX (AEOM-) ABOMICA INC.

XX Shannon ME, JI Y;

XX WPI; 2002-684026/74.

XX P-PSDB; ABG94709.

PT Novel GTP-Rho binding protein 2 and nucleic acids encoding the protein,

PT useful for the manufacture of a medicament for treating a disease

XX associated with altered expression or activity of human GRBP2 protein.

PS Claim 1; Page 48-49; 101pp; English.

XX The invention relates to an isolated GTP-Rho binding protein 2 (GRBP2)
polypeptide or a fragment of at least 6 amino acids or a sequence in
which at least 95% of deviations from GRBP2 sequences are conservative
substitutions. Also included are an isolated nucleic acid (GRBP2 NA)
encoding GRBP2 comprising the full length cDNA or CDS, fragments or
variants, GRBP2 vectors, host cells, antibodies, transgenic non-human
animals modified to contain GRBP2 NA (or unable to express the endogenous
orthologue of GRBP2), diagnosing a disease caused by a mutation in human
GRBP2 or altered expression of GRBP2, ant-agonists of GRBP2, GRBP2
microarrays, fusion proteins and screening for agents that modulate the
expression of GRBP2 NA. GRBP2 is useful for identifying binding partners
of GRBP2. GRBP2, GRBP2 NA and Ab are useful in therapy and in the
manufacture of a medicament for the treatment or prevention of a disorder
associated with increased or decreased expression or activity of human
GRBP2 (e.g. tumours, liposarcoma, ichthyosis congenita III and benign

DT	02-DEC-2002	(first entry)	Query Match	22.7%;	Score 493.2;	DB 6;	Length 3484;
XX			Best Local Similarity	56.7%;	Pred. No. 2e-84;		
DE		Human cDNA encoding GTP-Rho binding protein 2.	Matches	975;	Conservative	0;	Mismatches 733; Indels 12; Gaps 3;
XX							
KW	Human; ss;	GTP-Rho binding protein 2; GRBP2; chromosome 19q12; oncogene;	QY	219	GGGCTGTGACTCCCTGACGACGATCCAGTCGCGGCCAGCTGCGAGAGCCGCGAGGCCAGAT	278	
KW	tumour; liposarcoma; ichthyosis congenita III;		DB	89	GGGCTGTATATCCCTTGCACAAACCGCGCGAGTAAATTCAGAAATCAAAAGAGCTCTTT	148	
XX	benign familial infantile convulsion; gene therapy.						
OS	Homo sapiens.		QY	279	TCACGACGAGATTGACAGGAGCTGCAGATCGGAGCGGCGCTGAGAACTCTACAGAGC	338	
XX			DB	149	GAATCAGCAGATCTTGAAGCCGTCGGAAGAGGATCGAGCGGAAACCTCTTGAAGT	208	
PN	EP1231216-A2.		QY	339	CACGACAAACACCGGCTGAGAGACGCTCGCCCTGAGCTGAGCTAGCTCAACTCCAA	398	
XX	14-AUG-2002.		DB	209	GGCCACAAACTCAAAGGTGCGGAGCAAGTCGGCTGAGCTGAGCTTGTCACTCAG	268	
PF	17-JAN-2002; 2002EP-0001026.		QY	399	CCTGACGTGCTGAAAGAGGAGCTGAGAGCTCAGCGGTGCGTGGACCTTGGCGGCA	458	
XX	30-JAN-2001; 2001WO-US000663.		DB	269	CCTGACGATGCTCAAAGGAGAGCTGAGGGGTGAACATCTCGTGGCGTCTATCAGAA	328	
PR	30-JAN-2001; 2001WO-US000665.		QY	459	TGGGAGCGAAGCTGCTACTGTCCCATGATCCCTGATCCCTGATTCCTTGGCCCTGAAGGAAACGAAAGAGCT	518	
PR	30-JAN-2001; 2001WO-US000666.		DB	329	CACAGAGGAGGATTTACGATTCCTTGGCCCTGATTCCTTGGCCCTGAAGGAAACGAAAGAGCT	388	
PR	30-JAN-2001; 2001WO-US000667.		QY	519	GGACTGCTCTACACCGCTGAAGGAGCTGATCTCAGTGCACTTTGGAGAGGAGCGCGCTC	578	
PR	29-JUN-2001; 2001US-00895040.		DB	389	CGACTTTGCACTGCTCCTCAAGGATTTTATCCTGGAACATACACGTGAAGTGGCTATTT	448	
PA	(AEOM-) AEOMICA INC.		QY	579	CTACGAGGAGAAATCAGGAGCTGAGAGCCCTGCGGAGGCGCATGCGGACCCCGCGG	638	
PI	Shannon ME, JI Y;		DB	449	ATATGAAGATGAAATTCAGATCTTATGATCTGAGACAAAGCTTGTGCGAGCCCTAGCG	508	
XX	WPI; 2002-684026/74.		QY	639	GAATGAGTCGGGCTGAGAGCTGCTCACAGCCCTATTACACAGCTGTGCTTCTCGATGC	698	
DR	P-PSDB; ABG94709.		DB	509	GGATGAGCCGGGTGGAACCTGCTGATGACATCTTCATCAGCTGGGCTTTGTGAGAG	568	
XX			QY	699	CGCTTCTCTACCCCTGCCAGGAGCTCGGGCTCTTCTTCCACTGGTACGACTCGCTTAC	758	
XX			DB	569	TCGATTTCTCCGGCCACACGCGCAGATGGGACTCTCTTCCACTGGTATGACTCTCTAC	628	
CC			QY	759	TGGGTCCTCCGCGCCAGCAGCTGCGCTTGGCCCTTGGAGGCGGAGCGTCTCTTCAAAT	818	
CC			DB	629	CGGGTTCCTCCGCTCAGCCAGCAGAACTCTGCTGAGAGGAGCCAGTGTCTCTTCAAC	688	
CC			QY	819	CGGTGCTCTCACACGAGATTTGGGCGCGCAGACCGCTCTCTGACCGAGGGTGGCGG	878	
CC			DB	689	TGGGCGCTCTTACACCCAGATTTGGGACCCGCTGCGATCGGAGAGCGAGGCTGGGCTGA	748	
CC			QY	879	CCGCGCTATGAGGCTTTCCAGAGGCGCTTGGGCGCTTCCAGCCCTCTGAGGAGAACTT	938	
CC			DB	749	GAGTGCCATAGATGCTTTCCAGAGCGCGCAGGGGTTTAAATTAACCTGAAGACACATT	808	
CC			QY	939	CTCCATGCGCGGAGCCAGACATGAGGCTGCGCTTGGCCCTTGGCGCATGAGGAGCTCAT	998	
CC			DB	809	TACCATATCCAAATGACGACATGAGCCCTGCGCATGCTGAGCGTGTCTGCTCAAAATGAT	868	
CC			QY	999	GATGGCCAGGCGCAGGAAATGTGTGTTGAGGGCTCTCAACCTGCTCTCCATGGGCCCC	1058	
CC			DB	869	GCTTGACAGCCCAAGAGCGTGTGAGAAATCAGCCTTCTCTG-----GGATCCG	922	
CC			QY	1059	CCAAGACTGCTGGCCAGCTGCGCTTGGCGCAGGAGCGCCCGCAGGTGGGAGCGGAGTA	1118	
CC			DB	923	GAATGAATTTCTCATGCTGTGTAAGGTGGCTCAGGAGGCTCTAAGGTGGGAGGCTCTA	982	
CC			QY	1119	CAGGCTAGTGCACCGGACCATGCGCCAGCCAGCCGCTCCAGCTCCAGCTAGCTGTCTCTG	1178	
CC			DB	983	CCAACTACACGACGACCATGAGCCAGGCGCGGTGAAGAGAAATCCCTACTCTCTCT	1042	
CC			QY	1179	GATGCGCTCTGTCATGTCAAGCCGAGTACTTCCGCTCCCTGCGCCACTACACCTAGC	1238	
CC			DB	1043	GGCCAGTTAGCTGTGTAAGGCCCAACCACTAGCGGCGCTTGGCCCTACTTCTCTCTG	1102	
XX			QY	1239	CATGGCGCTCTGCGACGGCTCTCCCGCAGCGACCGGAGGAGAGCTCCCCACGACGAGGAGT	1298	

Novel GTP-Rho binding protein 2 and nucleic acids encoding the protein, useful for the manufacture of a medicament for treating a disease associated with altered expression or activity of human GRBP2 protein.

Claim 1; Fig 3; 101pp; English.

The invention relates to an isolated GTP-Rho binding protein 2 (GRBP2) polypeptide or a fragment of at least 6 amino acids or a sequence in which at least 95% of deviations from GRBP2 sequences are conservative substitutions. Also included are an isolated nucleic acid (GRBP2 NA) encoding GRBP2 comprising the full length cDNA or cDS, fragments or variants, GRBP2 vectors, host cells, antibodies, transgenic non-human animals modified to contain GRBP2 NA (or unable to express the endogenous orthologue of GRBP2), diagnosing a disease caused by a mutation in human GRBP2 or altered expression of GRBP2, ant-agonists of GRBP2, GRBP2 microarrays, fusion proteins and screening for agents that modulate the expression of GRBP2 NA. GRBP2 is useful for identifying binding partners of GRBP2. GRBP2 NA and Ab are useful in therapy and in the manufacture of a medicament for the treatment or prevention of a disorder associated with increased or decreased expression or activity of human GRBP2 (e.g. tumours, liposarcoma, ichthyosis congenita III and benign familial infantile convulsion, all associated with the chromosomal location of GRBP2, 19q12). GRBP2 is useful as a standard in immunoassay specific for the proteins, to be used in a therapeutic agent, as vaccines, to be and as antigens (e.g. for epitope mapping) or immunogens (e.g. for raising antibodies). GRBP2 NA is useful as hybridisation probes, to prime synthesis of nucleic acids, to prime first strand cDNA sequence, on an mRNA template, and to drive in vivo expression of the proteins. The vector is useful for shuttling GRBP2 NA between host cells derived from disparate organisms, for inserting GRBP2 NA into host cell chromosome, for expressing sense or antisense RNA transcripts of GRBP2 NA in vitro or within a host cell, and for expressing GRBP2 alone or as fusions to heterologous polypeptides. The antibody is useful as an analytical reagent for detection and quantification of GRBP2 and as an immuno therapeutic agent and is useful for flow cytometric detection, for scanning laser cytometric detection, or for fluorescent detection. The present sequence is a GRBP2 cDNA sequence

Sequence 3484 BP; 932 A; 814 C; 842 G; 896 T; 0 U; 0 Other;

Db 1103 CATCTCTCTCATCGACACACAGGTGAAGCAGGACGATCTGAGACACAGGAGAGTG 1162
QY 1299 CTTCTCTGACGAC-----CCCCACCTCTCTAAGCCCGAGGCCCTGTGC--TGCCGACAGGA 1352
Db 1163 CTTCTCTGACGACCTTAGACACATGCGAGGGGCTGACACCCCTTGCCACACCTGAGAA 1222
QY 1353 GCTGAGGAGCGCAGCAGCTTGGCAAGGACACCTGAGAGCGTGCATCTCTGGGAGGA 1412
Db 1223 TGATCAGCAGCGCGCAGCAGCTGAGGAGTCCCACTTGGCAGAGCCATGCTCATCAGCA 1282
QY 1413 GAGGCGCTGCGCTGACCGCTCTGCGCGCTGCGGAGGTGAGCTGCTTGGGC 1472
Db 1283 GGAGTCGGTGGGAGGAGAGCTCTGCAAGAGCTGCGGAGCATGAGTGCTACAGAA 1342
QY 1473 TGTGATCTCCAGACGCTGAGCGCTCACTGGCCAAAGTATGCGAGCTCGACCTGAGGA 1532
Db 1343 GGTGCTGTGTCGCGACAGGAACGCTCCCGGCTCAGTACGCCAGCACCAGGAGGAGGA 1402
QY 1533 TGAATCTGTGAGGCTGCGAGGCGCCCGGACATCCAGCCTTAAGACCCACAGAGCCAGA 1592
Db 1403 TGAATCTGTGAGGCTGCGAGGCGCCCGGACATCCAGCCTTAAGACCCACAGAGGTTGA 1462
QY 1593 GGCAGAGTCCAGCGCTGCTCCAGGAGGAGGCGCTGACATCTTCCATCGCTTGGGCGC 1652
Db 1463 CATATATATGCCCCAGTTCTCCAGCTGACAGTACAGGACTTCTTCCAGAGCTGGGCC 1522
QY 1653 CCTGTCTGTGTTCTCAGCAAGAACCGGTGGCGCTGGTGGGCGCCGTCCACCTGACCCG 1712
Db 1523 CTTATCTGTGTTTGGCTAAACAGCGGTGGAGCGCTCTCGAAGCATCCGCTTCACTGC 1582
QY 1713 AGGAGAGGCGGCTTGGCTCAGCTTGGGAGACTCGCTGCTCATCGCTGCGCT 1772
Db 1583 AGAAGAGGAGGACTTGGGCTTCACTTGAGAGGAGAACGCCCGGCTTCAAGTTCACTTCT 1642
QY 1773 CATTCAGGAGGAGCGCGCGCTGGCTGAGGAGGCGGCTACATTTGTGTGCT 1832
Db 1643 GGAATCTTACTGCTCTGCTGCTGGTGGAGAGCGCGGAGGAGATTAATTTGCTCCAT 1702
QY 1833 GAATGGGAGCGCATGAGGCTGGTGGAGACACGCGGAGGTGGTGGAGGCTGAGAGGCTGC 1892
Db 1703 TCAGTTGTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 1762
QY 1893 GGGAGAGCGGCGCGCGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 1932
Db 1763 TGGCGAGGAGGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 1802

RESULT 11

AAF58361
ID AAF58361 standard; cDNA; 3019 BP.
XX AC
XX AAF58361;
XX AC
DT 19-APR-2001 (first entry)
XX AC
XX Human GTP-binding associated protein #61 coding sequence.
DE DE
KW Human; guanosine triphosphate binding associated protein; GTP; GBAP;
KW inflammation; AIDS; Addison's disease; anaemia; arteriosclerosis; asthma;
KW autoimmune disorder; hepatitis; multiple sclerosis; cancer; diabetes;
KW osteoporosis; psoriasis; ss.
XX OS
XX Homo sapiens.
XX PN
XX WO200105970-A2.
XX AC
XX 25-JAN-2001.
XX AC
PF 19-JUL-2000; 2000WO-US019698.
XX AC
XX 19-JUL-1999; 99US-0144595P.
PR 23-AUG-1999; 99US-0150460P.
PR AC

PR 15-OCT-1999; 99US-0159849P.
XX (INCY-) INCYTE GENOMICS INC.
XX AC
PI Yue H, Tang YT, Bandman O, Hillman JL, Lal P, Au-Young J;
PI Reddy R, Yang J, Baughn MR, Lu DAM, Azimzai Y, Patterson C;
XX AC
DR WPI; 2001-091972/10.
DR P-PSDB; AAB68561.
XX AC
PT New guanosine triphosphate-binding associated proteins (GBAP) and their
PT encoding nucleic acids, useful for treating and/or diagnosing diseases
PT associated with GBAP expression, such as cancer, diabetes and asthma.
XX AC
PS Claim 5; Page 229-230; 233pp; English.
XX AC
CC The present invention relates to novel human guanosine triphosphate (GTP)
CC -binding associated proteins (GBAPs; AAB68561-AAB68566) and their coding
CC sequences (AAF58361-AAF58366). The proteins and coding sequences of the
CC present invention are useful for treating a variety of disorders
CC including inflammation, AIDS, Addison's disease, anaemia,
CC arteriosclerosis, asthma, autoimmune disorders, Grave's disease,
CC hepatitis, multiple sclerosis, cancer, diabetes, osteoporosis and
CC psoriasis
XX AC
SQ Sequence 3019 BP; 800 A; 708 C; 762 G; 748 T; 0 U; 1 Other;

Query Match 22.5%; Score 487.8; DB 4; Length 3019;
Best Local Similarity 56.9%; Pred. No. 2.1e-83;
Matches 957; Conservative 0; Mismatches 713; Indels 12; Gaps 3;
QY 257 TGCAGAGCCGACGGCCAGATTACACAGCAGATTCACAGAGGAGCTGCAGATCGGACCG 316
Db 190 TGCAGATCAAGAGCTGCTTTGATCAGCAGATCCTGAAAGCGTGGATGAGGACCG 249
QY 317 GCGCTGAGAACCTTCTACAGAGCCACAGCAACAAACCGGTGAGAGACCGTGGCCCTGG 376
Db 250 GAGCGGAAAACTTCTGAAAGTGGCCACAAACTCAAAGGTGCGGAGCAAGTGGCGCTGG 309
QY 377 AGCTGAGTACGCTCACTCCAACTGCTGCTGAGAGGAGCTGCAGAGCTCAGCG 436
Db 310 AGCTGAGTTCGCTCACTCAGACCTGAGATGCTCAAGAGAGAGCTGAGAGGCTGACAA 369
QY 437 GTGCGTGGACCTTGGCGGCGATGGGAGCGAAGCTGTCTGTCTCCCATATCCCTCTGG 496
Db 370 TCTCGTGGGCTCTATCAGAACACAGAGGAGCATTTACGATTCCTCTGATCTCTTG 429
QY 497 GCTTGAAGGAGACCAAGAGCTGAGTGTCTACACGCTGAGAGGAGCTGATCTCAGTGC 556
Db 430 GCTTGAAGGAGAACGAAAGAGCTGACCTTTGAGTCTCTCAAGGATTTTATCTTGAAC 489
QY 557 ACTTTGGAGAGGAGCGGCTCTCTACGAGGAGAGAAATCAGGAGCTGCAGGCGCTGGCG 616
Db 490 ATTACAGTGAAGATGGCTATTATATGAGATGAATTTGAGATCTTATGATCTGAGAC 549
QY 617 AGGCCATCGGACCCCGAGCGGAGTGTGGGCTGTGAGCTGCTCAGAGCTTATAC 676
Db 550 AAGCTTGTGAGCGCTAGCGGAGTGGGCGGCTGAGGCGGCTGGAATCTGATGACATCTCA 609
QY 677 ACCAGCTGTGCTTCTGAGTGGGCTTCTTCCCTGCTGAGGAGCTGCGGCTCTTCT 736
Db 610 TCCAGCTGGGCTTGTGAGAGTGTCTTCTTCCCGCCACACGCGAGATGAGATCTCTGT 669
QY 737 TCCACTGTGAGTCTGCTTACTGGGCTTCCCGGCGCCAGAGCTGCTGCTTGGAGTTCGAGA 796
Db 670 TCACCTGGTATGACTCTCTACCGGGTTCGGTTCAGCAGCAGAGAACCTGTGCTGGAGA 729
QY 797 AGGCGAGGCTTCTCTTCAACATCGGTGCTTCAACAGCAGATTTGGGCGCGCCAGGACC 856
Db 730 AGGCGAGGCTTCTTCAACATCGGTGCTTCAACAGCAGATTTGGGCGCGCCAGGACC 789
QY 857 GCTCTGACAGGAGGCTGCGCGCTATGGAGGCTTCCAGAGGCGCTGCGGCGCT 916

Db 790 GGCAGACGCGCTGGCTGGAGTGCATAGATGCTTTTCAGAGCCGCGAGGGTTT 849
QY 917 TCAGCCTCTGAGGAGAACTTCTCCATGCGCGAGCCAGACATGAGCGCTCGCTCCC 976
Db 850 TAAATTACCTGAAAGACATTTACCCATCTCAAAGTTACGACATGAGCCCTGCCATGC 909
QY 977 TCTGCCACTGGAGCAGCTCATGATGCCCGCCAGGCCAGGATGTGTTGAGGGCTCT 1036
Db 910 TCAGCGTCTGTCGCAAAATGATCTTGCAAGCCCAAGAAAGCGTGTGTTGAGAAATCA 969
QY 1037 CACCACTGCTCCATCGGCCCCCAAGACTGCTGCGCCAGCTGCGCTGGCCGAGAGG 1096
Db 970 GCCTTCTCTG-----GATCCNCAATGAATCTTCATGCTGTTGAAGTGGCTCAGAGG 1023
QY 1097 CGCCAGGTGGAGCGGAGTAGCAGGTAGTGACCGGACCATGCGCCAGCCACCCGCTCC 1156
Db 1024 CTGCTAAGGTGGAGAGGTCTACCAACAGCTACACGAGCCATGAGCCAGGCGCGGTGA 1083
QY 1157 ACGACTACGTGCTGCTCTCTGGAAGTCCCTGCTGCTGATGTCAGGCGGAGTACTTCGCT 1216
Db 1084 AAGAGAAATCCCTTACTCTCTGGCCAGCTTAGCTGCGTGAAGGCCACCACTACGCGG 1143
QY 1217 CCCTGGCCCACTACCACTAGCTAGCCCTCTGCGAGCGCTCCCGAGCGACCGAGGGAG 1276
Db 1144 CCCTGGCCCACTACTTCACTGCACTCTCTCATCGACCAAGTGAAGCCAGGACCGG 1203
QY 1277 AGCTCCCAACGACGAGAGGTCTTCTTGAGC-----CCCCACCTCTCTTAAGCCCGGA 1332
Db 1204 ATCTGACCAACGAGGAGAGTCTGCTCCAGCTCTACGACCACTGCGAGAGGGGCTGA 1263
QY 1333 GGCCTGTGC--TGCGGAGGAGCTGGAGGCGGAGGAGCTGCGGCTGACGCGCTGCTCGC 1450
Db 1264 CACCTTGGCCCACTGAAGAATGATCAGAGCGCGGAGAGTGGGGAAGTCCCACTTCG 1323
QY 1391 AGCTGCTCATCTTGGGCGAGGAGGCGCTGCGGCTGACGCGCTGCTGCGGCTCTCGC 1450
Db 1324 GCAGAGCCATGCTCATCAGGAGAGTCTGCTGGGAGGAGGAGCTCTGCAAGAGCTGC 1383
QY 1451 GCGAGTGACCTGCTTGGGCTGTGATCTCCAGACGCTGACGCGCTCACTGGCCAACT 1510
Db 1384 GGACGATTGAGGTGTACAGAGAGTGTGTGTGCGGACAGGAAAGTCTCCGCTCACT 1443
QY 1511 ATGCGGAGCTCACCGTGAAGTACTTCTGTGAGGCTGCGGAGGCGCGGACATCCAGC 1570
Db 1444 AGCCCAAGCAGGAGGAGTACTCTGTAACCTGATCAGCGCCCCCAGTGTGTG 1503
QY 1571 CTAAACCCACAGAGCCAGAGGCGAGATGCCAGCTGTCCAGGGGAAAGGGCGCTG 1630
Db 1504 CTAAACCTGAGCAAGAGGTGACATTAATTGCCCGAGTTCTCCAAGCTGACGTACGG 1563
QY 1631 ACATCTTCATCGGCTGGGCGCCCTGTCTGTGTTCTCAGCCAGAAACCGGTGGCGGCTGG 1690
Db 1564 ACTTCTTCCAGAGCTGGGCGCCCTTATCTGTGTTTTCGGCTAACCAAGCGGTGGAGCCTC 1623
QY 1691 TGGGCGCGTCCACCTGACCCAGAGAGGCGGCTTGGCCCTCAGCTTCCGGGAGACT 1750
Db 1624 CTCGAGATCCGCTTCACTGAGAGAGGAGGACTTGGGGTTCACCTTGAGAGGAAAG 1683
QY 1751 CGCTGTCTCATCGCTGCGCTCATTCAGGAGGAGCCAGCGCGGCTGCGCTGAGG 1810
Db 1684 CCCCCGTTTCAGTTTCACTTCTGGATCTTACTGTCTGCTCTGCTGCGTGGCAGAGCCCGG 1743
QY 1811 AGGGGACTACATTTGTGTGAGTGAATGGGACGCCATGAGGTGGTGGAGACACGCGGAGG 1870
Db 1744 AAGGAGATTATATTGCTCTCACTTCACTGATTTGATGTTGATGAGTGGCTGAGCTGAGT 1803
QY 1871 TGGTACGAGCTGAGGCTGCGGAGGAGGCGGCGCGGAGGCTGAGGTGTTGCTGCTGC 1930
Db 1804 TTATGAGCTGCTGAGAGCTTTGGCGAGGACGAGATCGAGATGAAGTCTGAGCCTCC 1863
QY 1931 TG 1932
Db 1864 TG 1865

RESULT 12

ABX97182
ID ABX97182 standard; cDNA; 2310 BP.
XX AC ABX97182;
XX DT 20-MAY-2003 (first entry)
XX DE Human NOV126a cDNA.
XX KW NOVX; cytosolic; cardiant; antiarteriosclerotic; antiasthmatic; cancer;
KW hypotensive; cardiomyopathy; bronchial asthma; gene therapy; vaccine;
KW human; gene; ss.
OS Homo sapiens.
XX PN WO200272757-A2.
XX PD 19-SEP-2002.
XX PF 08-MAR-2002; 2002WO-US006908.
XX PR 08-MAR-2001; 2001US-0274101P.
PR 08-MAR-2001; 2001US-0274194P.
PR 08-MAR-2001; 2001US-0274281P.
PR 08-MAR-2001; 2001US-0274322P.
PR 09-MAR-2001; 2001US-0274849P.
PR 12-MAR-2001; 2001US-0275235P.
PR 13-MAR-2001; 2001US-0275578P.
PR 13-MAR-2001; 2001US-0275579P.
PR 13-MAR-2001; 2001US-0275601P.
PR 14-MAR-2001; 2001US-0276000P.
PR 16-MAR-2001; 2001US-0276776P.
PR 19-MAR-2001; 2001US-0276994P.
PR 20-MAR-2001; 2001US-0277239P.
PR 20-MAR-2001; 2001US-0277321P.
PR 21-MAR-2001; 2001US-0277327P.
PR 22-MAR-2001; 2001US-0277791P.
PR 22-MAR-2001; 2001US-0277833P.
PR 23-MAR-2001; 2001US-0278152P.
PR 26-MAR-2001; 2001US-0278894P.
PR 27-MAR-2001; 2001US-0278999P.
PR 27-MAR-2001; 2001US-0279036P.
PR 28-MAR-2001; 2001US-0279344P.
PR 30-MAR-2001; 2001US-0277338P.
PR 30-MAR-2001; 2001US-0279995P.
PR 30-MAR-2001; 2001US-0280233P.
PR 02-APR-2001; 2001US-0280802P.
PR 02-APR-2001; 2001US-0280822P.
PR 02-APR-2001; 2001US-0280900P.
PR 04-APR-2001; 2001US-0281194P.
PR 13-APR-2001; 2001US-0283675P.
PR 30-APR-2001; 2001US-0287424P.
PR 02-MAY-2001; 2001US-0288066P.
PR 03-MAY-2001; 2001US-0288342P.
PR 15-MAY-2001; 2001US-0288528P.
PR 16-MAY-2001; 2001US-0291190P.
PR 16-MAY-2001; 2001US-0291099P.
PR 30-MAY-2001; 2001US-0294485P.
PR 31-MAY-2001; 2001US-0294889P.
PR 31-MAY-2001; 2001US-0294899P.
PR 18-JUN-2001; 2001US-0299027P.
PR 19-JUN-2001; 2001US-0299303P.
PR 19-JUN-2001; 2001US-0299310P.
PR 10-JUL-2001; 2001US-0304354P.
PR 31-JUL-2001; 2001US-0309198P.
PR 16-AUG-2001; 2001US-0312903P.
PR 10-SEP-2001; 2001US-0318462P.
PR 12-SEP-2001; 2001US-0318770P.
PR 27-SEP-2001; 2001US-0325430P.

PR 27-SEP-2001; 2001US-0325681P.
 PR 18-OCT-2001; 2001US-0330380P.
 PR 31-OCT-2001; 2001US-0335301P.
 PR 14-NOV-2001; 2001US-0332172P.
 PR 14-NOV-2001; 2001US-0332271P.
 PR 14-NOV-2001; 2001US-0332272P.
 PR 14-NOV-2001; 2001US-0333184P.
 PR 14-NOV-2001; 2001US-0333272P.
 PR 21-NOV-2001; 2001US-0332094P.
 PR 03-DEC-2001; 2001US-0337426P.
 PR 03-DEC-2001; 2001US-0338092P.
 PR 04-DEC-2001; 2001US-0337185P.
 PR 03-JAN-2002; 2001US-0345705P.
 PR 07-MAR-2002; 2002US-00092900.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 XX Padigar M, Spytek KA, Shenoy SG, Taupier RJ, Pena CE, Li L;
 PI Zerhusen BD, Gusev V, Ji W, Gorman L, Miller CE, Kekuda R;
 PI Patturajan M, Gangolli E, Vernet CAM, Guo X, Tchernev V;
 PI Fernandes ER, Casman SJ, Malyankar UM, Gerlach V, Liu Y, Anderson D;
 PI Spaderna SK, Catterton E, Burgess C, Leite M, Zhong H, Alsobrook JP;
 PI Lepley DM, Rieger DK;
 XX
 DR WPI; 2002-723332/78.
 DR P-PSDB; ABU65215.
 XX
 XX NOVX polypeptides and polynucleotides, useful for preventing or treating
 PT a disorder associated with aberrant NOVX expression or activity e.g.,
 PT cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial
 PT asthma.
 XX
 PS Claim 13; Page 470; 1103pp; English.
 XX
 CC This invention describes novel human NOVX polypeptides which have
 CC cytosolic, cardiac, antiarteriosclerotic, antiasthmatic and hypotensive
 CC activity. Pharmaceutical compositions comprising the NOVX proteins or
 CC nucleic acid molecules or NOVX antibodies are useful for preventing or
 CC treating a disorder associated with aberrant NOVX expression or activity
 CC e.g. cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial
 CC asthma. The products of the invention can be used for gene therapy or in
 CC a vaccine. ABX97008-ABX97185 are cDNA fragments amplified and isolated by
 CC the PCR primers and probes represented in ABX13460-ABX13462 and ABX97186-
 CC ABX97593. ABX97008-ABX97185 encode the NOVX proteins described in
 CC ABU65041-ABU65218
 XX
 SQ Sequence 2310 BP; 569 A; 622 C; 628 G; 491 T; 0 U; 0 Other;
 Query Match 21.4%; Score 463.6; DB 6; Length 2310;
 Best Local Similarity 56.7%; Pred. No. 8.1e-79;
 Matches 977; Conservative 0; Mismatches 729; Indels 16; Gaps 6;
 QY 219 GGGCTGTACTCCCTGACGAGATCCAGTGGCGGAGTGGAGCGCGAGGCCAGAT 278
 DB 114 GGGCTGTAAATCCCTTGCACAAACCGCGGAGTAAATTCAGAAATCAAGAGCTGCTTT 173
 QY 279 TCACGACAGATTGACAGAGTGCAGATCGGACGGCGCTGAGAACCTCTACAGAGC 338
 DB 174 GAATCAGAGATCTCTGAAGCCGTCGGATGAGGACCGGAGCGGAAACCTTCTGAAGT 233
 QY 339 CACCAGCAACACCGGGTGAGAGAGACGGTCCCTCTGGAGCTGAGCTACGTCACCTCCAA 398
 DB 234 GGCCACAAACTCAAGGTGCGGAGCAGTGGCTGGAGCTGAGCTTCGTCAACTCAGA 293
 QY 399 CCTGCAGCTGTGAAGGAGGAGCTGAGAGAGCTAGCGGTGGCGGTGGACCCCTGSCCGGCA 458
 DB 294 CCTGCAGATGCTCAAGGAGAGAGCTGGAGGGGCTGAACATCTCGGTGGCGGCTCTATCAGAA 353
 QY 459 TGGGAGCGAAGCTGTCAGTGTCCCATGATCCCTGGCGCTGAAGGAGAGCAAGAGAGCT 518
 DB 354 CACAGAGAGGAGATTACGATTCCTCCCTGATTCCTTGGCCCTGAAGGNAACGAAGAGCT 413
 QY 519 GGACTGTGTTACACCGCTGAAGAGAGCTGATCTCAGTGCACTTTGGAGAGGAGCGGCGCTC 578

Db 414 CGACTTTGTCAGTGTCTCTCAAGGATTTTATCTTGAACATTTACAGTCAAGATGGCTATTT 473
 QY 579 CTAAGAGGAGAAATCAGGAGCTGGAGGCTTGGCGAGGCCATGGGACCCCGCCAGCCG 638
 Db 474 ATATGAAGATGAATTTGCAGATCTTATGGATCTGAGACAGCTTTGTGGAGCGCTAGCCG 533
 QY 639 GAATGAGTGGGCTGGAGCTGTCTACAGCTTATTACAGCTTATTACAGCTGTGTCTCTGGATGC 698
 Db 534 GGATGAGGCGGGGTGAACTGCTGATGACAFACCTTATCAGCTGGCTTTGTTCGAGAG 593
 QY 699 GCGCTTCTCACCCTTGGCCAGGAGCTTGGGCTCTTTTCCACATGTTGACATCGCTTAC 758
 Db 594 TCGATTCTTCCCGCCACACGCGAGATGGGACTCTCTTCACTGGTATGACTCTCTCAC 653
 QY 759 TGGGGTCCCGCCAGCAGCGTGCCTTGGCCCTTCGAGAGGAGGAGCGCTTCTTCAACAT 818
 Db 654 CGGGGTTCCGGTCAGCCAGCAGAACCTGTCTGTCTGGAGAGGCGAGTGTCTGTTCACAC 713
 QY 819 CGGTGCGCTCCACACGAGATTTGGGCGCGCCAGGACCGCTCTCTGCACCGAGGAGTGC 878
 Db 714 TGGGGCTCTCTACCCAGATTTGGGACCGGTGTGATCGGACGAGCGAGGCTGGCTGGA 773
 QY 879 CCGCGCTATGAGGCGCTTCCAGAGGCGCTTGGGCGCTTCCAGCTCTCTGAGGAGAACTT 938
 Db 774 GAGTGCATAGATGCTCTTCCAGAGAGCGCGAGGGGTTTAAATTACCTGAAAGACACAT 833
 QY 939 CTCCATGCGCGAGCCAGCAGATGAGCGCTGCTCTCTGCGCAGCTGAGCAGAGCTCAT 998
 Db 834 TACCCATACCTCCAAAGTTACACATGAGCCCTGCCATGCTCAGCGTGTCTGTCAAATGAT 893
 QY 999 GATGGCCCGAGCCCGAGGAATGTGTGTTGAGGGGCTCTCACACAGCTGCTCTTCCATGGCCCC 1058
 Db 894 GCTTGCACAGCCCAAGAAAGCGTGTGAGAAATCAGCCTTCTCTG-----GGATCCG 947
 QY 1059 CCAAGACTGCTGGCCAGCTGCGCTTGGCGGAGGAGCGCCAGGTGGCAGCCGAGTA 1118
 Db 948 GAATGAATTTCTTATGCTGTGAAGTGGCTCAGGAGGCTGCTAAGTGGGAGAGGCTA 1007
 QY 1119 CAGGCTAGTGCACCGGACCATGGCCAGCCACCGCTCCAGACTACGCTGCTGCTCTCTG 1178
 Db 1008 CCAAGACTACAGCAGCCATGAGCAGGCGCGGTGAAGAGAAATCCCTACTCTCTG 1067
 QY 1179 GACTGCTCTGGTGCATGCAAGCCGAGTACTTCCGCTCTCTGCGCCACATACCACTGAGC 1238
 Db 1068 GGCAGCTTAGCTGCTGAGGCGCCACACTAGCGGCGCTGCGCCACTACTTCACTGTC 1127
 QY 1239 CATGGCCCTCTGGAGCGCTCCCGAGCGAGGAGAGCTCCCGACGACGAGCAGGT 1298
 Db 1128 CATCTCTCTCATCGACCACCAGGTGAAGCCAGGACCGGATCTGGACCCAGGAGAGATG 1187
 QY 1299 CTTCTCTCAGC---CCCCCACCCTCTCTAAGCCCGAGGCGCTGTGC--TGCCGCGAGGA 1352
 Db 1188 CCGTCCCGAGCTTACGACACATGCGAGAGGGCTGACACCTTGGCCACACTGAAGAA 1247
 QY 1353 GCTGGAGAGCGCAGGAGCTTGGCAAGGCACACTGAAGCGTGCCTTCTGGGGCAGGA 1412
 Db 1248 TGATCAGCAGCGCGCAGCTGGGGAAGTCCCACTTGGCGAGAGCCATGGCTCATCAGGA 1307
 QY 1413 GGAGCGCTGGGCTGACGCGCTGTGCGCGGTCTGCGGAGGTGGACCTGCTTCCGGC 1472
 Db 1308 GGAGTGGGTGGGAGGCGCGCTCTGCAAGAGCTCGGAGCATTGAGGTGCTTACAGAA 1367
 QY 1473 TGTGATCTCCAGACGCTGACAGCGCTCACTGGCCCAAGTATGCGGAGGTTCGACCGTGA 1532
 Db 1368 GGTGCTGTGTCGCGCAGGNAACGCTCCCGGCTCACGTCACGCCACCGAGGAGGAGGA 1427
 QY 1533 TGACTTCTGTAGAGTGCAGGAGGCC---GGACATCCAGCTTAAGACCCACCAAGAGCCA 1590
 Db 1428 TGACCTGTGAACCTGATTCGACGCGCCCGAGAGTGTGTTGTTAAACTGAGCAAGAGGTT 1487
 QY 1591 GAGCCGAGGATGCCACGCTGTCCGAGGGAAGGGCTGACATCTTCCATCGGCTGGG 1650

1488 GACATTATTTGCCCCAGTTCT-CCAGCTGACAGTCAAGGACTTCTTCAGAGCT-GGG 1545
1651 CCCTCTCTGTCTTCTCAGCAAGAACCGGTGGGCTGTGGGCGCCGTCACCTGACC 1710
1546 CCCTTATCTGTCTGTCTGCGCTAACAGCGGTGACGCTCTCTCGAAGCATCCGTTCACT 1605
1711 CGAGGAGAGGGCGGCTTTGGCCCTCAGCTTTCGGGGAGACTCGCTCTCTCATCGTGGC 1770
1606 GCAGAGAGAGGGAGCTTGGGTTTCACTTGAAGAGGAACGCCCGTTAGGTTCACTTC 1665
1771 GTCAATTCAGGGAGCGAGCGCGGCTTGAAGAGGGCGACTACATTTGTCTCA 1830
1666 CTGGATCTTACTGTCTCTGCTCGGTGGCAGGAGCCGGAGAGGATATATTTCTTC 1725
1831 GTCAATGGCAGCCATGAGGTGTGGAGACACGCGAGGTGGTGAAGGCTTGAAGGCT 1890
1726 ATTACGTTGTGGATTGTAAGTGGCTGACGCTGAGTGAGGTTATGAAGCTGCTGAAGAGC 1785
1891 GCGGGAGAGCGGGCGCCAGCCTGACAGGTGGTGTCTGCTGCTG 1932
1786 TTGGCGAGGACGAGATCGAGATGAAGTCGTGAGCCTCTG 1827

RESULT 13
AAZ15546
ID AAZ15546 standard; cDNA; 764 BP.
XX AAZ15546;
AC AAZ15546;
XX AAZ15546;
DT 12-OCT-1999 (first entry)
XX Human gene expression product cDNA sequence SEQ ID NO:3015.
DE Human; gene; gene expression product; diagnosis; therapy; probe;
XX Human; gene; gene expression product; diagnosis; therapy; probe;
KW detection; mapping; tissue typing; profiling; forensic; cancer;
XX genetic analysis; colorectal cancer; breast cancer; lung cancer; ss.
XX Homo sapiens.
XX Homo sapiens.
XX WO9938972-A2.
XX 05-AUG-1999.
XX 28-JAN-1999; 99WO-US001619.
XX 28-JAN-1998; 98US-0072910P.
XX 24-FEB-1998; 98US-0075954P.
XX 31-MAR-1998; 98US-0080114P.
XX 03-APR-1998; 98US-0080515P.
XX 03-APR-1998; 98US-0080666P.
XX 21-OCT-1998; 98US-0105234P.
XX 28-OCT-1998; 98US-0105877P.
XX (CHIR) CHIRON CORP.
XX (HYSE-) HYSEQ INC.
XX
XX Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J;
XX Reinhard C, Giese K, Randazzo F, Kennedy GC, Pot D, Kassam A;
XX Lamson G, Drmanac R, Crkvenjakov R, Dickson M, Drmanac S, Labat I;
XX Leshkowitz D, Kita D, Garcia V, Jones WL, Stache-Crain B;
XX WPI; 1999-494092/41.
XX
XX Novel human genes and their expression products which are differentially
XX expressed in different cell types.
XX
XX Claim 1; Page 1452; 2479pp; English.

CC suspected of being cancerous, where the gene product is encoded by one of
CC the 5248 polynucleotide sequences given in AAZ12532 to AAZ17779. The
CC polynucleotides can be used as a source of primers and probes, which can
CC be used for a variety of purpose, e.g detection of expression levels,
CC mapping, tissue typing or profiling, forensics, genetic analysis and
CC detection of polymorphisms. Polypeptides encoded by the polynucleotides
CC can be used for raising antibodies for experimental, diagnostic and
CC therapeutic purposes. The polynucleotides may also be used to construct
CC arrays for diagnostics (which may be used to determine function of an
CC encoded protein); and to detect differences in expression levels between
CC two cells (e.g. to identify abnormal or diseased tissue in a human, to
CC identify a genetic predisposition or susceptibility to a disease such as
CC cancer). The polynucleotides of the invention are especially used in the
CC diagnosis, prognosis and management of colorectal cancer, breast cancer,
CC and lung cancer. The polynucleotides can also be used to screen for
CC peptide analogues and antagonists
XX

Seq Sequence 764 BP; 157 A; 224 C; 224 G; 137 T; 0 U; 22 Other;

Query Match 20.1%; Score 436; DB 2; Length 764;
Best Local Similarity 99.1%; Pred. No. 1.3e-73;
Matches 447; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1719 GGGCGGCTTTGGCCCTCAGCTTTCGGGGAGACTGCGCTGCTCATCGTCGCTCATTC 1778
DB 51 GGGCGGCTTTGGCCCTCAGCTTTCGGGGAGACTGCGCTGCTCATCGTCGCTCATTC 110
QY 1779 AGGGAGCCAGGCGCGGGCGGCTGGAAGAGGGGCGACTACATTTGTCTGCTGATGG 1838
DB 111 AGGGAGCCAGGCGCGGGCGGCTGGAAGAGGGGCGACTACATTTGTCTGCTGATGG 170
QY 1839 GCAGCCATGCAAGTGGTGGAGACACGCGGAGGTGGTGACGGAGCTGAAGGCTCGGGAGA 1898
DB 171 GCAGCCATGCAAGTGGTGGAGACACGCGGAGGTGGTGACGGAGCTGAAGGCTCGGGAGA 230
QY 1899 GGGCGGGCGGCGGCTGCGAGGTGGTGTGCTGCTGCCAGCTTAGACTGCCAGCTTGGG 1958
DB 231 GGGCGGGCGGCGGCTGCGAGGTGGTGTGCTGCTGCCAGCTTAGACTGCCAGCTTGGG 290
QY 1959 GGACCGCGCGGCGGCTGCTGCGGGCGGCGGCTTCTAAGGAGCCAGAGGAGCATGG 2018
DB 291 GGACCGCGCGGCGGCTGCTGCGGGCGGCGGCTTCTAAGGAGCCAGAGGAGCATGG 350
QY 2019 TTGCAAGACCCCGGCGCATCCAGTGGGGCCAGTCCCCGGGCGGCTTCACTGAGCGGAAA 2078
DB 351 TTGCAAGACCCCGGCGCATCCAGTGGGGCCAGTCCCCGGGCGGCTTCTAAGGAGCCAGG 410
QY 2079 GGCCGAGGAGGCAAGACTGAGAGGTGCGCCCGAGCCCTGTGCCCCAGTGAAGCCAGCTCC 2138
DB 411 GGCCGAGGAGGCAAGACTGAGAGGTGCGCCCGAGCCCTGTGCCCCAGTGAAGCCAGCTCC 470
QY 2139 GCGCTCATCTTTGAAGCACCCAGGCTGGCCG 2169
DB 471 G-CCTCATCTTTGAAGCACCCAGGCTGGCCG 500

RESULT 14
AAC91349

ID AAC91349 standard; cDNA; 2757 BP.

XX AAC91349;
AC AAC91349;
XX AAC91349;
XX AAC91349;
DT 16-MAR-2001 (first entry)
XX Human polynucleotide for diagnostics and therapeutics, SEQ ID NO: 49.

XX Human; diagnostics and therapeutics; dithp; cytostatic;
XX immunosuppressive; nootropic; neuroprotective; antiviral; antiallergic;
XX hepatotropic; antidiabetic; antiinflammatory; antitumor; vulnary;
XX anticonvulsant; antibacterial; antifungal; antiparasitic; cardiant;
XX cancer; immune disorder; cardiovascular disorder; neurological disease;
XX infection; endocrine disorder; metabolic disorder; ss.
XX

Db 1075 GAGCCTCTGCAAGAGCTCCG-GANATTGAGTGCTTACAGAGGTGTGTGTCG-CACA 1132
QY 1491 GCAGCGCTCACTGCCAAGTATGCGAGCTCGACCGTGAGGATGACTTCTGTGAGGCTGC 1550
Db 1133 GGAACGCTCCCGGCTACGCTACGCCAGCCAGAGGAGGATGACCTGTGTAACCTGAT 1192
QY 1551 CGAGGCGCCGACATCCAGCCTTAAGACCCACAGAGCCAGAGCCAGAGTGCACGCCT 1610
Db 1193 CGAGCGCCCGAGTGTGTCTGCTAAACTGAGCAAGAGTTGACATTATATGCG-CCATT 1251
QY 1611 GTCCAGGGGAAGGGGCTGACATCTTCCATCGGCTGGGGCCCTGTCTGTCTTCAGC 1670
Db 1252 CTCGAAGCTGACAGTCAAGGACTTCTTCCAGAGCTGGGCGCTTATCTGTGTTCGGC 1311
QY 1671 CAAAGACCGTGGCGGTGTGGGGCCCGTCCACTGACCCGAGGAGGCGGCTTTGG 1730
Db 1312 TAAAGAGCGGTGACGCTCTCTCGAAGCATCCGCTTCACTGCAGAGAGAGGAGCTTGG 1371
QY 1731 CCTCAGCTTCGGGGAGACTCGCTCTCTCATCGCTGCGCTCATTCAGGAGGAGCCAGC 1790
Db 1372 GTTCACTTGAGAGGAGACGCCCGCTTCAAGTTCAGGTTCACTTCTCGATCCTTACTGCTGC 1431
QY 1791 CGCGCGGCTGGCTGAAGAGGCGGACTACATTGTGTGATGATGGGAGCCATGCGAG 1850
Db 1432 CTCGCTGGCAGGAGCCCGGAGGAGATTATTTGTCTCCATTCACTGTGTGATTGTA 1491
QY 1851 GTGCTGGACACGCGGAGGTGTGACGAGCTGAAGGCTGCGGGAGAGG 1900
Db 1492 GTGGCTGACGCTGAGTGAGGTTATGAAGCTGCTGAAGAGCTTGGCGAGG 1541

RESULT 15

ABK92255

ID ABK92255 standard; DNA; 1671 BP.

XX AC ABK92255;

XX 15-AUG-2002 (first entry)

XX Prostate cancer-associated DNA sequence #141.

XX Prostate cancer; prostate tumour tissue; human; mammal; cytostatic;

XX gene therapy; gene; ds.

XX Mammalia.

XX PN WO200230269-A2.

XX PD 18-APR-2002.

XX PF 12-OCT-2001; 2001WO-US032045.

XX 13-OCT-2000; 2000US-00687576.

XX PR 08-DEC-2000; 2000US-00733288.

XX PR 08-DEC-2000; 2000US-00733742.

XX PR 24-JAN-2001; 2001US-0263957P.

XX PR 16-MAR-2001; 2001US-0276791P.

XX PR 16-MAR-2001; 2001US-0276888P.

XX PR 06-APR-2001; 2001US-0281922P.

XX PR 30-APR-2001; 2001US-0286214P.

XX PR 04-MAY-2001; 2001US-0288589P.

XX PA (EOSB-) EOS BIOTECHNOLOGY INC.

XX PI Gish KC, Mack DH, Wilson KE, Afar D, Hevezi P;

XX DR P-PSDB; ABG61936.

XX DR WPI; 2002-471335/50.

XX DR P-PSDB; ABG61936.

XX PT Detecting a prostate cancer-associated transcript in a cell in a patient,
PT useful for diagnosing prostate cancer (PC) or screening modulators of PC,
PT by determining if prostate cancer-associated genes are expressed in a

PT prostate tissue.

XX Claim 22; Page 417-418; 436pp; English.

XX The present invention relates to methods of detecting a prostate cancer-
XX associated transcript in a cell from a patient. The method comprises
XX contacting a biological sample from the patient with prostate cancer-
XX associated polynucleotides (designated PC genes) that selectively
XX hybridise to a sequence that is at least 80% identical to them. The
XX prostate cancer-associated polynucleotide sequences are differentially
XX expressed in prostate tumour tissue or in prostate cancer and are derived
XX from the tissues of various organisms such as humans or other mammals
XX (e.g. mice, sheep and dogs). The methods of the invention are useful for
XX diagnosing and treating prostate cancer in mammals. The prostate cancer-
XX associated genes are useful for diagnosing or treating prostate cancer,
XX as well as for identifying modulators of prostate cancer or agents that
XX inhibit prostate cancer. The nucleic acid sequences are particularly
XX useful in gene therapy, as a vaccine or in antisense applications.
XX ABK92115-ABK92263 represent prostate cancer-associated polynucleotide
XX sequences

SQ Sequence 1671 BP; 402 A; 459 C; 457 G; 353 T; 0 U; 0 Other;

Query Match 17.5%; Score 379.2; DB 6; Length 1671;

Best Local Similarity 56.0%; Pred. No. 8.7e-63;

Matches 783; Conservative 0; Mismatches 603; Indels 12; Gaps 3;

QY 541 GAGCTGATCTCAGTGCACCTTTGAGAGGAGCGGCGCTCTACGAGCGAATFACGGAG 600

Db 1 GATTTTATCTCGAACAATTACAGTGAAGATGGCTATTTATATGAAGATGAATTCAGAT 60

QY 601 CTGGAGGCGCTCGCGCAGGCCATGCGGACCCCGCAGCGGAATCAGTCGGCGCTGGAGCTG 660

Db 61 CTTATGGATCTGAGACAAGCTTGTGCGACCGCTAGCCGGGATGAGCCGGGGTGAAGTG 120

QY 661 CTCACAGCCTATTACAACAGCTGTGCTTCTCGATGCGCGCTTCTCAACCTTCCAG 720

Db 121 CTGATGACATACTTCATCCAGCTGGGCTTTGTGAGAGTCTGATTTCCGCGCCACACG 180

QY 721 AGCCTGGGCTCTTCTTCCACTGCTGAGCTCGCTTCTGAGTGGGCTCCCGGCCAGCGCT 780

Db 181 CAGATGGGACTCCTGTTTCACTGGTATGACTCTCTCACCGGGGTTCCGGTCCAGCAG 240

QY 781 GGCCTGGGCTTCGAGAAGGGCAGCTTCTTCTCAACATCGGTGCGCTCCACAGCAGATT 840

Db 241 AACCTGTCTGTGAGAGAGCCAGTGTCTTGTTCACACTGCGGGGCTCTACACCCAGATT 300

QY 841 GGGGCGCGCAGGACCGCTTCTGCAACGAGGTGCGCGCGCTATGAGAGCTTCCAG 900

Db 301 GGGACCCGCTGTGATCGGCAGACGAGCTGGGCTGGAGAGTGCCTATAGTGCCTTTTCA 360

QY 901 AGGGCGCTGGGCGCTTTCAGCCTCTGAGGAGAACTTCTCCATGCGCGCAGCCAGAC 960

Db 361 AGAGCCGCGAGGGGTTTAAATTTACCTGAAGACACATTTACCCATACTCCAGTTTACGAC 420

QY 961 ATGAGCGCTGCGTCCCTCTGCGCACTGGAGCAGCTCATGATGGCCCGCAGCCCGAATGT 1020

Db 421 ATGAGCGCTGCGTCCCTCTGCGCACTGGAGCAGCTCATGATGGCCCGCAGCCCGAATGT 480

QY 1021 GTGTTTGAAGGCGCTTCAACACCTGCTCCATGCGCCCGCCCAAGACTGCTGGCCAGCTG 1080

Db 481 GTGTTTGAAGGCGCTTCAACACCTGCTCCATGCGCCCGCCCAAGACTGCTGGCCAGCTG 534

QY 1081 CGCCTGGGCGAGGAGCGCGCCAGTGGCAGCCAGTACAGGCTAGTGCAACCGGACCATG 1140

Db 535 AAGGTGGCTCAGGAGGCTGTAAAGTGGGAGAGGCTTACCAACAGCTACACGAGCCATG 594

QY 1141 GCCCAGGCCACCGCTCCAGACTAGTGTGCTCTCTCTGAGCTGCGCTGGTGTGATGTCAAG 1200

Db 595 AGCCAGGCGCGGTGAAGAGAACATCCCTTACTTCTGGGCCAGCTTAGCCTGGGTGAAG 654

QY 1201 GCCGAGTACTTCCGCTCCCTGGGCCACTACACAGTACGATGGCCCTCTGCGAGGCTCC 1260

Db 655 GCCCACCACCTACCGGCCCTGGCCCACTTCTCCTGCCATCCTCTCATCGACACCAG 714
Qy 1261 CCAGCCAGCGAGGAGAGTCCCAACGACGAGAGGTCTTCTCTGAGC----CCCCAC 1316
Db 715 GTGAAGCCAGGACCGGATCTGGACCAACAGGAGAGTGCCTGTCCAGCTCTACGACCAC 774
Qy 1317 CTCTCTTAAGCCCGGAGGCCCTGTGCT--GCCCAGGAGCTGGAGGAGCGCAGGACGCTT 1374
Db 775 ATGCCAGAGGGGTGACACCCCTTGGCCACACTGAAGAAATGATCAGCAGCGCCGACAGCTG 834
Qy 1375 GGCAAGGCACACCTGAAGCGTGCATCCTGGGGCAGGAGGCGGTGCGGCTGCACGCC 1434
Db 835 GGAAGTCCCACTTGGCAGAGCCATGGCTCATCAAGAGAGTGGTGGCGGAGGCCAGC 894
Qy 1435 CTGTGCGCGCTCTGGCGAGGTGGACCTCTCTGGGCTGTGATCTCCAGACGCTGCAG 1494
Db 895 CTCTGAAGAAGCTGGCGAGCATTTAGGTGCTACAGAAGTGTGTGTGCCGCACAGGAA 954
Qy 1495 CGTCACTGCCAAGTATGGGAGCTCGACCGTGGAGTACTTCTGTGAGGCTGCCGAG 1554
Db 955 CGTCCCGGCTCAGTACGCCCCAGCACAGGAGGAGTACCTGTGAACCTGATCGAC 1014
Qy 1555 GCCCGGACATCCAGCCTAAGACCCACCAAGCCAGAGCCAGGATGCCACGCTGTCC 1614
Db 1015 GCCCCAGTGTGTGTCTAAACTGAGCAAGAGGTTGACATTATTTGCCCACTTCTCC 1074
Qy 1615 CAGGGGAAGGGGCTGACATCTTCATCGGCTGGGGCCCTGTCTGTGTCTCAGCCAAG 1674
Db 1075 AAGCTGACAGTCACGGACTTCTCCAGAAGCTGGGGCCCTTATCTGTGTTTTCGGCTAAC 1134
Qy 1675 AACGGTGGCGCTGGTGGGGCCCTCCACTGACCCGAGGAGGGCGGCTTTGGGCTC 1734
Db 1135 AAGCGTGGACGCTCTCGAAGCATCCGCTTCACTGCAAGAAAGGGGACTTGGGGTTC 1194
Qy 1735 ACGCTTCGGGAGACTCGCCTGTCTCATCGCTGCCCTCAFTCCAGGAGCCAGGCCCG 1794
Db 1195 ACCTTGAGAGGGAACGCCCGCTTCAGGTTCACTTCTCGGATCCTTACTGTCTGCGCTCG 1254
Qy 1795 GCGCTGCGCTGAAGAGGGGCGACTACATTTGTGTGATGAATGGGCGAGCCATGCGGTGG 1854
Db 1255 GTGGCAGGAGCCCGGAAGGAGATATATTTGTCTCCATTCACTTGTGGATTGTAAGTGG 1314
Qy 1855 TGGAGACACGGGAGGTGGTACCGAGCTGAAGGCTGGGAGAGGGCGGCCAGCCCTG 1914
Db 1315 CTGACGCTGAGTGAGGTATGAAGCTCTGAAGAGCTTTGGCGAGGACGATCGAGATG 1374
Qy 1915 CAGGTGGTGTGCGCTGCTG 1932
Db 1375 AAAGTCTGTAGCCTCTG 1392

Search completed: July 13, 2004, 02:51:08
Job time : 564.787 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 12, 2004, 20:08:35 ; Search time 5564.67 Seconds
(without alignments)
16894.276 Million cell updates/sec

Title: US-10-697-266-1_COPY_103_2271

Perfect score: 2169

Sequence: 1 atgatctctggaggagagcc.....tgaagcaccacgggtggccg 2169

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sy.*

13: gb.un.*

14: gb.vi.*

15: em.ba.*

16: em.fun.*

17: em.hum.*

18: em.in.*

19: em.mu.*

20: em.om.*

21: em.or.*

22: em.ov.*

23: em.pat.*

24: em.ph.*

25: em.pl.*

26: em.ro.*

27: em.sts.*

28: em.un.*

29: em.vi.*

30: em.htg.hum.*

31: em.htg.inv.*

32: em.htg.other.*

33: em.htg.mus.*

34: em.htg.pln.*

35: em.htg.rod.*

36: em.htg.mam.*

37: em.htg.vrt.*

38: em.sy.*

39: em.htgo.hum.*

40: em.htgo.mus.*

41: em.htgo.other.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2169	100.0	2469	6	AR269115	Sequence
2	2169	100.0	2469	6	AX642309	Sequence
3	1851	89.9	2837	6	AX675063	Sequence
4	1951	89.9	3763	9	BC025767	Homo sapi
5	1866	86.0	2088	9	AY082588	Homo sapi
6	1865	86.0	3647	9	AB067516	Homo sapi
7	1102.2	50.8	1932	10	MMU43194	Mus muscu
8	1036.6	47.8	2301	10	BC052010	Mus muscu
9	798.8	36.8	4729	6	BD183400	Novel gen
10	558.8	25.8	3230	4	CFA347749	Canis fam
11	496.4	22.9	3516	9	BC036447	Homo sapi
12	494.8	22.8	2061	9	HSB347750	Homo sapi
13	494.8	22.8	3500	6	AX835273	Sequence
14	494.8	22.8	3500	9	AK098246	Homo sapi
15	487.8	22.5	3019	6	AX077672	Sequence
16	485.6	22.4	3397	9	HSB03254	Sequence
17	477.4	22.0	2233	5	BC044556	Danio rer
18	475.8	21.9	2224	5	AY391456	Homo sapi
19	443.4	20.4	2262	9	AF268032	Homo sapi
20	436	20.1	764	6	BD214873	Novel hum
21	418.2	19.3	2757	6	AX593072	Sequence
22	311	14.3	3798	9	AK126506	Homo sapi
23	272.8	12.6	2403	3	AF132025	Drosophil
24	222.4	10.3	178343	2	AC024049	Homo sapi
25	217	10.0	19025	6	AR269116	Sequence
26	217	10.0	19025	6	AX642311	Sequence
27	217	10.0	179057	2	AF276758	Sequence
28	217	10.0	179941	9	AC105118	Homo sapi
29	168.2	7.8	61188	2	AC135076	Homo sapi
30	165.4	7.6	123016	2	AF189001	Homo sapi
31	165.4	7.6	124060	9	AC087793	Homo sapi
32	165	7.6	165	6	BD213706	Novel hum
33	149.6	6.9	2300	6	AX833345	Sequence
34	149.6	6.9	2300	9	AK095001	Homo sapi
35	146.6	6.8	232687	2	AC116393	Mus muscu
36	143.4	6.6	261608	2	AC097158	Rattus no
37	143.4	6.6	272301	2	AC133673	Rattus no
38	127	5.9	3174	2	AC014171	Drosophil
39	127	5.9	185200	3	AC094401	Drosophil
40	127	5.9	191558	3	AC013431	Drosophil
41	127	5.9	303345	3	AE003500	Drosophil
42	110	5.1	5234	9	AB025194	Homo sapi
43	108.4	5.0	2925	6	AX713572	Sequence
44	108.4	5.0	2925	9	AK055157	Homo sapi
45	108.4	5.0	4265	9	AB040904	Homo sapi

ALIGNMENTS

RESULT 1
AR269115
LOCUS AR269115
DEFINITION Sequence 1 from patent US 6500655.
ACCESSION AR269115
VERSION AR269115.1 GI:29699963
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2469)
AUTHORS Rusch,D., Ketchum,K.A., Di Francesco,V. and Beasley,E.M.
TITLE Isolated human kinase proteins, nucleic acid molecules encoding human kinase proteins, and uses thereof
JOURNAL Patent: US 6500655-A 1 31-DEC-2002;

AR269115 2469 bp DNA linear PAT 10-APR-2003

Db 1663 GACATCCAGCCTAAGACCCACACAGAACCCAGAGCCAGAGCCAGGATGCCAGCCTGTGCCAGGG 1722
QY 1621 AAGGGGCTTACATCTTCATCGGCTGGGGCCCTCTGTGTCTCAGCAAGAACCGG 1680
Db 1723 AAGGGGCTTACATCTTCATCGGCTGGGGCCCTCTGTGTCTCAGCAAGAACCGG 1782
QY 1681 TGGCGGCTGGTGGGGCCCTCCACCTGACCCGAGGAGAGGGCGCTTTGGCCTCAGCTT 1740
Db 1783 TGGCGGCTGGTGGGGCCCTCCACCTGACCCGAGGAGAGGGCGCTTTGGCCTCAGCTT 1842
QY 1741 CGGGAGAGCTCGCCTGTCTCATCTGCTGCGCTCAATCCAGGGAGCCAGGCGCGGGCT 1800
Db 1843 CGGGAGAGCTCGCCTGTCTCATCTGCTGCGCTCAATCCAGGGAGCCAGGCGCGGGCT 1902
QY 1801 GGCCTGAAGAGAGGGCGGCTACATTTGTCTAGTGAATGGGGAGCCATGCGAGTGGGAGA 1860
Db 1903 GGCCTGAAGAGAGGGCGGCTACATTTGTCTAGTGAATGGGGAGCCATGCGAGTGGGAGA 1962
QY 1861 CAGCGGAGGTGGTACGGAGCTGAGGCTGCGGGAGAGCGGGCGGCGGCTGCGAGGTG 1920
Db 1963 CAGCGGAGGTGGTACGGAGCTGAGGCTGCGGGAGAGCGGGCGGCGGCTGCGAGGTG 2022
QY 1921 GTGTGCTGTGCTGCCAGCTCTAGACTGCCAGCTTGGGGAGCCGCGGCGGCTGCTGCTG 1980
Db 2023 GTGTGCTGTGCTGCCAGCTCTAGACTGCCAGCTTGGGGAGCCGCGGCGGCTGCTGCTG 2082
QY 1981 GGGCCCAAGGGGCTTTAAGAGCCAGAGGAGCATGGTTGCAAGACCCCGGCATCCACG 2040
Db 2083 GGGCCCAAGGGGCTTTAAGAGCCAGAGGAGCATGGTTGCAAGACCCCGGCATCCACG 2142
QY 2041 TGGGCGAGTCCCGGCGGCTTCACTCACTGAGCCGAAAGGCCAGAGGCGCAAGACTGGA 2100
Db 2143 TGGGCGAGTCCCGGCGGCTTCACTCACTGAGCCGAAAGGCCAGAGGCGCAAGACTGGA 2202
QY 2101 GGTGCCCCCAGGCGCTGTGCCAGTGAAGCAGCTCCGCGCTCATCTTGAAGACCCA 2160
Db 2203 GGTGCCCCCAGGCGCTGTGCCAGTGAAGCAGCTCCGCGCTCATCTTGAAGACCCA 2262
QY 2161 GGGTGGCGG 2169
Db 2263 GGGTGGCGG 2271

RESULT 3
AX675063 2837 bp DNA linear PAT 27-MAR-2003
LOCUS Sequence 11 from Patent WO02077235.
DEFINITION AX675063
ACCESSION AX675063
VERSION AX675063.1 GI:29333362
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Burford,N., Ding,L., Yue,H., Thornton,M., Wallia,N.K., Gandhi,A.R.,
Arvizu,C., Baughn,M.R., Swarnakar,A., Duggan,B.M., Lu,D.A.,
Thangavelu,K., Warren,B.A., Tang,J.F., Khan,F.A., Yao,M.G. and
Emerling,B.M.
TITLE Intracellular signaling molecules
JOURNAL Patent: WO 02077235-A 11 03-OCT-2002;
Incyte Genomics, Inc. (US)
FEATURES Location/Qualifiers
source
1..2837
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 5560161CBI"

ORIGIN
Query Match 89.9%; Score 1951; DB 6; Length 2837;
Best Local Similarity 100.0%; Pred. No. 8.1e-295;
Matches 1951; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 219 GGGCTGTGACTCCCTGACGAGATCCAGTGCAGCCAGCTGCAGAGCCGAGGCGCCAGAT 278
Db 181 GGGCTGTGACTCCCTGACGAGATCCAGTGCAGCCAGCTGCAGAGCCGAGGCGCCAGAT 240
QY 279 TCACAGCAGATGACAAAGAGCTGCAGATGCGGAGCGGGCGCTGAGAACTCTACAGAGC 338
Db 241 TCACAGCAGATGACAAAGAGCTGCAGATGCGGAGCGGGCGCTGAGAACTCTACAGAGC 300
QY 339 CACACGACACACCGGCTGAGAGAGCGGTGCGCCCTGGAGCTGAGCTACGCTCAACTCCAA 398
Db 301 CACACGACACACCGGCTGAGAGAGCGGTGCGCCCTGGAGCTGAGCTACGCTCAACTCCAA 360
QY 399 CTTGAGCTGCTGAAGAGAGCTGCAGAGAGCTCAGCGGTGGCGTGGAGCCCTGGCGGCA 458
Db 361 CTTGAGCTGCTGAAGAGAGCTGCAGAGAGCTCAGCGGTGGCGTGGAGCCCTGGCGGCA 420
QY 459 TGGAGCGAAGCTGTCACTGTCCCATGATCCCGCTGGGCGCTGAAGAGACCAAGAGCT 518
Db 421 TGGAGCGAAGCTGTCACTGTCCCATGATCCCGCTGGGCGCTGAAGAGACCAAGAGCT 480
QY 519 GGAAGCTGTACACCGCTGAAGAGCTGATCTCAGTGCATCTTTGGAGAGACCGCGCTC 578
Db 481 GGAAGCTGTACACCGCTGAAGAGCTGATCTCAGTGCATCTTTGGAGAGACCGCGCTC 540
QY 579 CTACAGAGCAGAAATCAGGGAGCTGGAGGCCCTTGGCGAGGCCATGCGGACCCGAGCG 638
Db 541 CTACAGAGCAGAAATCAGGGAGCTGGAGGCCCTTGGCGAGGCCATGCGGACCCGAGCG 600
QY 639 GAATGAGTCCGGCTGGAGCTGCTCAGAGCTTATACACAGCTGTGCTTCTGAGATG 698
Db 601 GAATGAGTCCGGCTGGAGCTGCTCAGAGCTTATACACAGCTGTGCTTCTGAGATG 660
QY 699 GCGCTTCTCACCCCTGCCAGAGCTTGGGCTCTTCTTCCACTGTGTACGCTTAC 758
Db 661 GCGCTTCTCACCCCTGCCAGAGCTTGGGCTCTTCTTCCACTGTGTACGCTTAC 720
QY 759 TGGGGTCCCGGCCAGCAGCTGCTTGGGCTTTCGAGAGGCGAGCTTCTTCAACAT 818
Db 721 TGGGGTCCCGGCCAGCAGCTGCTTGGGCTTTCGAGAGGCGAGCTTCTTCAACAT 780
QY 819 CGGTGCGCTCCACGAGATTTGGGCGCGCCAGGAGCGCTTCTGACCGAGGCTGCGCG 878
Db 781 CGGTGCGCTCCACGAGATTTGGGCGCGCCAGGAGCGCTTCTGACCGAGGCTGCGCG 840
QY 879 CCGGCTATGAGAGCCTTCCAGAGGCGCTTGGGCGCTTCCAGCTTCTGAGGGAGAACTT 938
Db 841 CCGGCTATGAGAGCCTTCCAGAGGCGCTTGGGCGCTTCCAGCTTCTGAGGGAGAACTT 900
QY 939 CTCCCATGCGCGAGCCAGACATGAGGGCTGCGCTTCCCTCTGCGCACTGGAGAGCTCAT 998
Db 901 CTCCCATGCGCGAGCCAGACATGAGGGCTGCGCTTCCCTCTGCGCACTGGAGAGCTCAT 960
QY 999 GATGGCCAGGCGCCAGGAATGTTGAGGGCTTTCAGGCGCTTCTCACCTGCTCCATGGCCCC 1058
Db 961 GATGGCCAGGCGCCAGGAATGTTGTTGAGGGCTTTCAGGCGCTTCTCACCTGCTCCATGGCCCC 1020
QY 1059 CCAAGACTGCTGCGCCAGCTGCGCTTGGCGCAGAGGCGCCCGAGGTGGAGCGAGTA 1118
Db 1021 CCAAGACTGCTGCGCCAGCTGCGCTTGGCGCAGAGGCGCCCGAGGTGGAGCGAGTA 1080
QY 1119 CAGGCTAGTGACCGGACCATGGGCGCAGCCACCGTCCACAGCTACGCTGCTCCTG 1178
Db 1081 CAGGCTAGTGACCGGACCATGGGCGCAGCCACCGTCCACAGCTACGCTGCTCCTG 1140
QY 1179 GACTGCGCTGTGATGCAAGGCGAGTACTTCCGCTCCCTGGGCCCACTACCGAGTAC 1238
Db 1141 GACTGCGCTGTGATGCAAGGCGAGTACTTCCGCTCCCTGGGCCCACTACCGAGTAC 1200
QY 1239 CATGGCCCTCTGCGAGCGCTCCCGAGCGACCGAGGAGAGCTTCCCGACGCGAGCAGGT 1298
Db 1201 CATGGCCCTCTGCGAGCGCTCCCGAGCGACCGAGGAGAGCTTCCCGACGCGAGCAGGT 1260


```

/misc_feature
/db_xref="AAH25767.1"
/db_xref="GI:19343951"
/db_xref="LocusID:114822"
/translation="MILERPDCAGAGEBSPRLQGDLSITQCGQLSRRRAIHQOI
DKELMRTCAENLYRATSNRRVETVALELSVYVNSLQLLKEELELSGVDPPGRHGS
EAVTVPMLPLGKEYKELDWSTPLKELISVHFEDEGASVEARIRELEALROAMRTPSR
NESGLELUTAYNQJCLFDARFLTPARSLGLFPHWYDLSITGVPAQORALAPEKGSVLIF
NITGALHTQAGQDSCTEGARRAMEAFORAGAFSLLREAFSPDMSAASLAL
EQLMMAQAQCEVEGLSPSPAPQDLAQLRAQAAQVAABEYRHLVHTMAQPPVHD
YVPSWALVHVHKAERYFLAHVHMAALCDGSPATEGELPHEQVFLQPPTSSRPRG
PVLPOELERROLGKAHLRAILIOBEALRLHALCLREVLLRAVLSQTLQSLAK
YABLREDDFCAAEPADIQPKTHQPEARMPLRSQKGPDIPLRGLPLSVFSKRW
RLVGPVHLTRGEGGFLTGKDSVPLIAAVIPGSQAAAGLKEGDIVSVNGPCRW
RHAEEVTEUKAAGEAGSLQVSLPSLSDRRPVLPRGLLRSLQREHGCKTP
ASTWASPRLLNWSRAQQGTGCGPQPCAPVKPAPPSSLKHPGW"
204..341
/notes="H31; Region: H31 repeat"
/db_xref="CDD:pfam02185"
423..881
/notes="BRO1; Region: BRO1-like domain. This functionally
uncharacterised domain is found in a number of signal
transduction proteins, including RhoPhilin and BRO1"
/db_xref="CDD:pfam03097"
1641..1841
/notes="PDZ; Region: Domain present in PSD-95, Dlg, and
ZO-1/2. Also called DHR (Dlg homologous region) or GLGF
(relatively well conserved tetrapeptide in these domains).
Some PDZs have been shown to bind C-terminal polypeptides"
/db_xref="CDD:smart00228"

```

ORIGIN

```

Query Match      89.9%; Score 1951; DB 9; Length 3763;
Best Local Similarity 100.0%; Pred. No. 7.5e-295;
Matches 1951; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 219 GGGGTGACTCCCTGACGAGATCCAGTGGCGGACGTGCAGAGCCGCGAGGCCAGAT 278
DB 161 GGGGTGACTCCCTGACGAGATCCAGTGGCGGACGTGCAGAGCCGCGAGGCCAGAT 220
QY 279 TCACAGAGATTGACAGAGAGCTGCAGATCGGACCGGGCGCTGAGAACCTTACAGAGC 338
DB 221 TCACAGAGATTGACAGAGAGCTGCAGATCGGACCGGGCGCTGAGAACCTTACAGAGC 280
QY 339 CACAGCAACAAACGGGTGAGAGAGACGGTGCCTCGAGTGCAGTGCAGTGCAGTGCAGT 398
DB 281 CACAGCAACAAACGGGTGAGAGAGACGGTGCCTCGAGTGCAGTGCAGTGCAGTGCAGT 340
QY 399 CTGCGAGTGTGAGAGAGAGCTGCAGAGAGCTGCAGAGTGGCGGCTGAGAACCTTACAGAGC 458
DB 341 CTGCGAGTGTGAGAGAGAGCTGCAGAGAGCTGCAGAGTGGCGGCTGAGAACCTTACAGAGC 400
QY 459 TGGAGAGCAAGCTGTCACTGTCCCATGATCCCTCGGCGCTGAGAGAGCAAGAGAGCT 518
DB 401 TGGAGAGCAAGCTGTCACTGTCCCATGATCCCTCGGCGCTGAGAGAGCAAGAGAGCT 460
QY 519 GGAAGTGTGTAACCGCTGAGAGAGCTGATCTCAGTGCAGTTCGAGAGAGCGGCGCTC 578
DB 461 GGAAGTGTGTAACCGCTGAGAGAGCTGATCTCAGTGCAGTTCGAGAGAGCGGCGCTC 520
QY 579 CTACAGAGCAGAAATCAGGAGCTGGAGCCCTCGGCGAGGCCATCGCGACCCCGAGCGG 638
DB 521 CTACAGAGCAGAAATCAGGAGCTGGAGCCCTCGGCGAGGCCATCGCGACCCCGAGCGG 580
QY 639 GAATGAGTCGGGCTTGGAGCTGCTCAGAGCTATTACACAGAGCTGTGCTTCTGATGC 698
DB 581 GAATGAGTCGGGCTTGGAGCTGCTCAGAGCTATTACACAGAGCTGTGCTTCTGATGC 640
QY 699 GGGCTTCTTCAACCCCTGCCAGAGCTCGGGCTCTTCTTCCACTGTAAGAGCTCGCTTAC 758
DB 641 GGGCTTCTTCAACCCCTGCCAGAGCTCGGGCTCTTCTTCCACTGTAAGAGCTCGCTTAC 700
QY 759 TGGGGTCCCGGCCAGAGAGCTGCGCTTCCAGAGAGAGGAGCGGCTTCTTCAACAT 818

```

```

DB 701 TGGGGTCCCGGCCAGCAGCGTGCCTTCCAGAGAGGCGAGGGTCTCTTCAACAT 760
QY 819 CGGTGCTCCACACGAGATTGGGGCGCGCCAGAGACCGCTCTTGACAGAGGGTGCCTCG 878
DB 761 CGGTGCTCCACACGAGATTGGGGCGCGCCAGAGACCGCTCTTGACAGAGGGTGCCTCG 820
QY 879 CGCGCTATGAGAGGCTTCCAGAGGGCGCTGGGGCTTTCAGCTCTTGGAGGAGACTT 938
DB 821 CGCGCTATGAGAGGCTTCCAGAGGGCGCTGGGGCTTTCAGCTCTTGGAGGAGACTT 880
QY 939 CTCCTATGCGCGAGCCCGAGACATGAGCGCTGGCTCCCTTGGCGACTGGAGCAGCTCAT 998
DB 881 CTCCTATGCGCGAGCCCGAGACATGAGCGCTGGCTCCCTTGGCGACTGGAGCAGCTCAT 940
QY 999 GATGGCCAGGCGCCAGGAATGTGTGTTGAGGGCTCTTCCACACCTGCCTTCCATGCCCC 1058
DB 941 GATGGCCAGGCGCCAGGAATGTGTGTTGAGGGCTCTTCCACACCTGCCTTCCATGCCCC 1000
QY 1059 CCAAGACTGCTGGCCAGCTGCGCTGGCGAGAGGGCGCCAGGTGGCAGCGAGTA 1118
DB 1001 CCAAGACTGCTGGCCAGCTGCGCTGGCGAGAGGGCGCCAGGTGGCAGCGAGTA 1060
QY 1119 CAGGCTAGTGACCGGACCATGCGCCAGCCAGCCCGTCCACGACTAGTGTGCTCTCTCTG 1178
DB 1061 CAGGCTAGTGACCGGACCATGCGCCAGCCAGCCCGTCCACGACTAGTGTGCTCTCTCTG 1120
QY 1179 GACTGCGCTGCTGATGTCAAGCCCGAGTACTTCCGCTCCCTGGCGCCACTTACACGTA 1238
DB 1121 GACTGCGCTGCTGATGTCAAGCCCGAGTACTTCCGCTCCCTGGCGCCACTTACACGTA 1180
QY 1239 CATGGCCCTCTGCGAGCGGCTCCCGAGCGCCAGGAGAGCTCCCGCCAGCGAGGAGGT 1298
DB 1181 CATGGCCCTCTGCGAGCGGCTCCCGAGCGCCAGGAGAGCTCCCGCCAGCGAGGAGGT 1240
QY 1299 CTTCCTGACGCGCCCGACCTCTCTAAGCCCGCGAGGCGCTGTGTGCTGCGCGAGAGCTGGA 1358
DB 1241 CTTCCTGACGCGCCCGACCTCTCTAAGCCCGCGAGGCGCTGTGTGCTGCGCGAGAGCTGGA 1300
QY 1359 GGAGCGAGCGAGCTTGGCAAGGCACACCTGAGCGTGCCATCTCTGGGGCAGAGAGAGG 1418
DB 1301 GGAGCGAGCGAGCTTGGCAAGGCACACCTGAGCGTGCCATCTCTGGGGCAGAGAGAGG 1360
QY 1419 GCTGCGGCTGACGCGCTGTGCGCGCTGCTGCGCGAGGTGAGACCTGCTCGGGCTGTGAT 1478
DB 1361 GCTGCGGCTGACGCGCTGTGCGCGCTGCTGCGCGAGGTGAGACCTGCTCGGGCTGTGAT 1420
QY 1479 CTCCAGAGCGTGCAGCGCTCAGTGGCCAAAGTATGCGAGCTCGACCTGAGAGGATGAT 1538
DB 1421 CTCCAGAGCGTGCAGCGCTCAGTGGCCAAAGTATGCGAGCTCGACCTGAGAGGATGAT 1480
QY 1539 CTGTGAGGCTGCGCGAGGCGCCGAGCATCCAGCTTAAGACCCACAGAAAGCCAGAGCCAG 1598
DB 1481 CTGTGAGGCTGCGCGAGGCGCCGAGCATCCAGCTTAAGACCCACAGAAAGCCAGAGCCAG 1540
QY 1599 GATGCGAGCTGCTCCAGGGGAGGGGCTGAGCATCTTCCATCGGCTGGGCGCCCTGTC 1658
DB 1541 GATGCGAGCTGCTCCAGGGGAGGGGCTGAGCATCTTCCATCGGCTGGGCGCCCTGTC 1600
QY 1659 TGTGTTCTCAGCAAGAACCGTGGCGCTGGTGGGGCGCGCTTCCATCGGCTGGGCGCCCTGTC 1718
DB 1601 TGTGTTCTCAGCAAGAACCGTGGCGCTGGTGGGGCGCGCTTCCATCGGCTGGGCGCCCTGTC 1660
QY 1719 GGGCGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 1778
DB 1661 GGGCGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 1720
QY 1779 AGGAGAGCCAGGCGCGGCGGCTGGCTTGAAGAGGGCGGCTTGAAGAGGGCGGCTTGAAGAG 1838
DB 1721 AGGAGAGCCAGGCGCGGCGGCTGGCTTGAAGAGGGCGGCTTGAAGAGGGCGGCTTGAAGAG 1780
QY 1839 GCAGGCGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 1898
DB 1781 GCAGGCGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 1840

```


1020 GACTGCCCTGGTGTGATGTCAGAGCCGAGTACTTCCGCTCCCTGGCCCACTACACGCTAGC 1079
1239 CATGGCCCTCTGGCAGCGCTCC----- 1261
1080 CATGGCCCTCTGGCAGCGCTCCGCTGAGTGCCACACACTTGCCTCCATGCTACTGCCAAG 1139
1262 -----CAGCGACCGAGGAGCTCCC 1283
1140 GCCCCCCCGCGCAGGCTCACAGCTCTCTGTCCCCCAGCAGCGACCGAGGAGGAGCTCCC 1199
1284 CAGCAGCAGCAGCTTCTCTGAGCCGCCCACTCTCTTAAGCCGCCGAGGCGCTGTGCT 1343
1200 CAGCAGCAGCAGCTTCTCTGAGCCGCCCACTCTCTTAAGCCGCCGAGGCGCTGTGCT 1259
1344 GCCCGCAGGAGCTGGAGGAGCGCAGGAGCTGTGGCAGGCAACCTGAAGCGTGCATCCT 1403
1260 GCCCGCAGGAGCTGGAGGAGCGCAGGAGCTGTGGCAGGCAACCTGAAGCGTGCATCCT 1319
1404 GGGCAGGAGGAGCGCTGCGCTGCAGCCCTGTGCGCGTCTGCGCGAGGTGACCT 1463
1320 GGGCAGGAGGAGCGCTGCGCTGCAGCCCTGTGCGCGTCTGCGCGAGGTGACCT 1379
1464 GCTTGGCGCTGATCTCCAGACCGCTGCAGCGCTCACTGGCCCAAGTATCGGAGCTCGA 1523
1380 GCTTGGCGCTGATCTCCAGACCGCTGCAGCGCTCACTGGCCCAAGTATCGGAGCTCGA 1439
1524 CCGTGAGGATGACTTCTGTGAGGTGCGGAGCGCCCGGACATCCAGCTTAAGCCACCA 1583
1440 CCGTGAGGATGACTTCTGTGAGGTGCGGAGCGCCCGGACATCCAGCTTAAGCCACCA 1499
1584 GAAGCCAGGCGCAGGATGCCACCGCTGTCCAGGGAAGGGGCTGACATCTTCATCG 1643
1500 GAAGCCAGGCGCAGGATGCCACCGCTGTCCAGGGAAGGGGCTGACATCTTCATCG 1559
1644 GCTGGGCGCCCTGTCTGTGTTCTCAGCAAGAACCGGTGGCGGTGGGGCCCGTCCA 1703
1560 GCTGGGCGCCCTGTCTGTGTTCTCAGCAAGAACCGGTGGCGGTGGGGCCCGTCCA 1619
1704 CTTGACCCGAGGAGGCGGCTTGGCTCAGCTTCGCGGAGACTGCGCTGCTCAT 1763
1620 CTTGACCCGAGGAGGCGGCTTGGCTCAGCTTCGCGGAGACTGCGCTGCTCAT 1679
1764 CGCTGCGCTCATTTCCAGGAGCGCGCGCGCTGCGCTGAAGGAGGCGGACTACAT 1823
1680 CGCTGCGCTCATTTCCAGGAGCGCGCGCGCTGCGCTGAAGGAGGCGGACTACAT 1739
1824 TGTGTCACTGAATGGCAGCCATGAGGTGTGGAGACACCGGAGGTGTGACGAGCT 1883
1740 TGTGTCACTGAATGGCAGCCATGAGGTGTGGAGACACCGGAGGTGTGACGAGCT 1799
1884 GAAGGCTCGGAGAGGCGGCGCGCGCTGAGGTGTGCTGCTGCGCGCGGCTCTAG 1943
1800 GAAGGCTCGGAGAGGCGGCGCGCGCTGAGGTGTGCTGCTGCGCGCGGCTCTAG 1859
1944 ACTGCCAGCTTGGGGAGCGCGCGCGCTGCTGCTGCGCGCGCGCGCGCTTAAAGAG 2003
1860 ACTGCCAGCTTGGGGAGCGCGCGCGCTGCTGCTGCGCGCGCGCGCTTAAAGAG 1919
2004 CCAGAGGAGCATGTTGTCAGACCCCGGCACTCAGGTGGGCGAGTCCCGGCGCGCTCT 2063
1920 CCAGAGGAGCATGTTGTCAGACCCCGGCACTCAGGTGGGCGAGTCCCGGCGCGCTCT 1979
2064 CAACTGGAGCCGAAAGGCCAGGCGGAGAGTGGAGGCTGCCCCGAGCTGTGCGCC 2123
1980 CAACTGGAGCCGAAAGGCCAGGCGGAGAGTGGAGGCTGCCCCGAGCTGTGCGCC 2039
2124 AGTGAAGCCAGCTCGCCCTCATCTCTTGAAGCAACCCAGGAGTGGCGG 2169
2040 AGTGAAGCCAGCTCGCCCTCATCTCTTGAAGCAACCCAGGAGTGGCGG 2085

RESULT 6
AB067516

LOCUS AB067516 3647 bp mRNA linear PRI 15-SEP-2001
DEFINITION Homo sapiens mRNA for KIAA1929 protein, partial cds.
ACCESSION AB067516
VERSION AB067516.1 GI:15620916
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 Nagase, T., Kikuno, R. and Ohara, O.
AUTHORS Prediction of the coding sequences of unidentified human genes.
TITLE XXI. The complete sequences of 60 new cDNA clones from brain which
code for large proteins
JOURNAL DNA Res. 8 (4), 179-187 (2001)
MEDLINE 21456161
PUBMED 11572484
REFERENCE 2 (bases 1 to 3647)
AUTHORS Chara, O., Nagase, T. and Kikuno, R.
TITLE Direct Submission
JOURNAL Submitted (27-JUL-2001) Osamu Chara, Kazusa DNA Research Institute,
Department of Human Gene Research; 1532-3, Yana, Kisarazu, Chiba
292-0812, Japan (E-mail:cdnainfo@kazusa.or.jp.
URL:http://www.kazusa.or.jp/huge, Tel:81-438-52-3913,
Fax:81-438-52-3914)
FEATURES
Location/Qualifiers
1..3647
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="fj08985"
/tissue_type="brain"
/note="vector:pBluescriptII SK plus"
1..3647
/gene="KIAA1929"
1..1233
/gene="KIAA1929"
/note="Start codon is not identified. fj08985 cDNA clone
for KIAA1929 has 349-bp, 320-bp and 413-bp insertions
after the positions 450, 936 and 1283 of the sequence of
KIAA1929, respectively."
/codon_start=1
/product="KIAA1929 protein"
/protein_id="BAB67822.1"
/db_xref="GI:15620917"
/translation="EERP DGAGAEESPR LQCGDSL TQ I QCG LQSRRA QHQIDKE
LQMTGAENLYRATSNRVRVETVALSLSYVNSLQQLKELELGGVDPGRHSEAV
TVPMPLGKETKELMSTPLKELISVHFGEDGASYEAEIRLEALRQAMTPGRNES
GLELLTAYNQLCFELDARFLTPARSLGLFHHYDLSLTPAQQRALAFKGSVLPNIG
ALHTQIGARDQRCTEGARAMEAFQAGAFSLLENFSAFSPDMSAALCALEQL
MMAQACVCEGLSPSPAPQDCQLQLRAQEAQVAAYRLVHRTVAQPPVHDYVP
VSWTALVHVKAIEYFRSLAHVHVAVALCDGSRCPPLPMLVLPFRPRAGLTASLSPSSD
RGRAPHARAGLPAAPHL"

ORIGIN

Query Match 86.0%; Score 1865; DB 9; Length 3647;
Best Local Similarity 96.3%; Pred. No. 2e-281;
Matches 1951; Conservative 0; Mismatches 0; Indels 76; Gaps 1;
QY 219 GGGCTGTGACCTCCCTGACGCGAGATCCAGTGGCGGCGAGTGCAGAGCCGCGAGGCCCGCAGAT 278
DB 51 GGGCTGTGACCTCCCTGACGCGAGATCCAGTGGCGGCGAGTGCAGAGCCGCGAGGCCCGCAGAT 110
QY 279 TCACCAGCAGATTGCAAGGAGCTGCAGATCGGACCGGGCGCTGAGAACCTTCTACAGAGC 338
DB 111 TCACCAGCAGATTGCAAGGAGCTGCAGATCGGACCGGGCGCTGAGAACCTTCTACAGAGC 170
QY 339 CACCAGGAAACACCGGGTGAGAGACCGTCCCGCTGGAGCTGAGCTACCTCACTCCA 398
DB 171 CACCAGGAAACACCGGGTGAGAGACCGTCCCGCTGGAGCTGAGCTACCTCACTCCA 230
QY 399 CCTGCAGCTGCTGAAGGAGGAGCTGCAGGAGCTCAGCGGTGGCGTGGACCTTGGCGGCA 458

231 CCTGCAGCTGCTGAAGAGAGCTGGAGAGCTCAGCGGTGGCGTGGACCCCTGGCCGGCA 290
459 TGGGAGCGAGCTGTCACTTCCCAATGATCCCTTGGGCTGAAGAGAGCAAGAGCT 518
291 TGGGAGCGAGCTGTCACTTCCCAATGATCCCTTGGGCTGAAGAGAGCAAGAGCT 350
519 GGAAGCTGTACACCGCTGAAGAGCTGTCACTTCAAGTGTCACTTGGAGAGAGCGCGCTC 578
351 GGAAGCTGTACACCGCTGAAGAGCTGTCACTTCAAGTGTCACTTGGAGAGAGCGCGCTC 410
579 CTACGAGGAGCAAAATCAGGAGCTGGAGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 638
411 CTACGAGGAGCAAAATCAGGAGCTGGAGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 470
639 GAATGAGTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 698
471 GAATGAGTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 530
699 GGGCTTCCCTCAGCCCTGCGAGAGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 758
531 GGGCTTCCCTCAGCCCTGCGAGAGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 590
759 TGGGCTTCCCTCAGCCCTGCGAGAGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 818
591 TGGGCTTCCCTCAGCCCTGCGAGAGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 650
819 CGGTGCTTCCACACGAGATTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 878
651 CGGTGCTTCCACACGAGATTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 710
879 CGGTGCTTCCACACGAGATTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 938
711 CGGTGCTTCCACACGAGATTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 770
939 CTCCATGCGCGAGCCAGACATGAGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 998
771 CTCCATGCGCGAGCCAGACATGAGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 830
999 GATGCGCGAGCCAGGAGATTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 1058
831 GATGCGCGAGCCAGGAGATTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 890
1059 CCAAGACTGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 1118
891 CCAAGACTGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 950
1119 CAGGCTAGTGCACCGAGCCATGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 1178
951 CAGGCTAGTGCACCGAGCCATGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 1010
1179 GACTGCGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 1238
1011 GACTGCGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 1070
1239 CATGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 1261
1071 CATGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 1130
1262 -----CAGCGAGCCAGGAGAGCTTC 1282
1131 GCGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 1190
1283 CCAGCGAGGAGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 1342
1191 CCAGCGAGGAGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 1250
1343 TGCGCGAGGAGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 1402
1251 TGCGCGAGGAGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 1310
1403 TGGGCGAGGAGGAGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 1462
1311 TGGGCGAGGAGGAGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 1370

1463 TGCTTCGGGCTGTGATCTCCAGACCTGCAGCGCTCACTGGGCAAGTATGCGAGCTCG 1522
1371 TGCTTCGGGCTGTGATCTCCAGACCTGCAGCGCTCACTGGGCAAGTATGCGAGCTCG 1430
1523 ACCGTGAGGATGACTTCTGTGAGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 1582
1431 ACCGTGAGGATGACTTCTGTGAGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 1490
1583 AGAAGCCAGAGCGGAGGATGCGACGCTGTCCAGGGGAAAGGGCTTGAATCTTCAATC 1642
1491 AGAAGCCAGAGCGGAGGATGCGACGCTGTCCAGGGGAAAGGGCTTGAATCTTCAATC 1550
1643 GGTGGGGGCTTGTCTGTGTTCTCAGCCAAAGACCGGTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 1702
1551 GGTGGGGGCTTGTCTGTGTTCTCAGCCAAAGACCGGTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 1610
1703 ACTGACCCGAGGAGGAGGCTTGTGGCTTCACTGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 1762
1611 ACTGACCCGAGGAGGAGGCTTGTGGCTTCACTGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 1670
1763 TCGCTGCGCTTCACTTCCAGGAGCGAGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 1822
1671 TCGCTGCGCTTCACTTCCAGGAGCGAGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 1730
1823 TTGTGTCACTGATTAATGGGAGAGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 1882
1731 TTGTGTCACTGATTAATGGGAGAGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 1790
1883 TGAAGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 1942
1791 TGAAGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 1850
1943 GACTGCGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 2002
1851 GACTGCGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 1910
2003 GCCAGAGGAGCATGTTTGAAGACCCCGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 2062
1911 GCCAGAGGAGCATGTTTGAAGACCCCGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 1970
2063 TCAACTGAGCGGAGGAGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 2122
1971 TCAACTGAGCGGAGGAGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 2030
2123 CAGTGAAGCAGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 2169
2031 CAGTGAAGCAGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 2077

RESULT 7

MMU43194
LOCUS Mus musculus rhophilin mRNA, complete cds. linear ROD 02-FEB-1996
DEFINITION U43194
ACCESSION U43194.1 GI:1176421
VERSION rho binding protein; PKN homology.
KEYWORDS Mus musculus (house mouse)
SOURCE ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1932)
AUTHORS Watanabe, G., Saito, Y., Madaule, P., Ishizaki, T., Fujisawa, K., Morii, N., Mukai, H., Ono, Y., Kakizuka, A. and Narumiya, S.
TITLE Protein kinase N (PKN) and PKN-related protein rhophilin as targets of small GTPase Rho
JOURNAL Science 271 (5249), 645-648 (1996)
MEDLINE 96165390
PUBMED 8571126
REFERENCE 2 (bases 1 to 1932)
AUTHORS Watanabe, G. and Narumiya, S.
TITLE Direct Submission
JOURNAL Submitted (13-DEC-1995) Go Watanabe, Pharmacology, Kyoto University

QY	639	GAATGAGTGGGCTGGAGTGTCTACAGCCTATTACACACAGCTGTGCTTCTGATGC	698
Db	1111		1111
QY	566	GGACGAGGGGGCTGGATCTGCTCGCAGCTACTATAGCAGCTCTGTGTTCTGACGC	625
Db	1111		1111
QY	699	CGCTTTCTTACCCCTGCGAGAGCTCTGGGCTTTCTTCCACTGTGTACACTCGCTTAC	758
Db	1111		1111
QY	626	ACGTTTCTTCACTTCCAGGAGCCCGGGCTGTCTTCCACTGTGTACACTCGCTAC	685
Db	1111		1111
QY	759	TGGGCTCCCGCCAGCAGCGTGCCTTGGCTTTTCGAGAGGGAGCGTCTTCTTCAACAT	818
Db	1111		1111
QY	686	GGGGTCCCGCAGCAGCGGGCTCTGGCTTTTCGAGAGGGAGTGTACTGTTTAACAT	745
Db	1111		1111
QY	819	CGTGGCTTCCACACGAGATTGGGGCGCGCAGGACCGTCTCTGACACAGGGTGCCTG	878
Db	1111		1111
QY	746	TGGGGCACTCCACACGAGATTGGAGCACGCGAGGATTGTCTCTGCAACCGAGGACAA	805
Db	1111		1111
QY	879	CGCGCTATGAGGCTTCCAGAGGGCGCTGGGGCTTCAGCTCTCTGAGGGAGAACTT	938
Db	1111		1111
QY	806	CCAGCTGTGAGCTTTCCAGAGGGCTGTGGGCTTCAGACTCTCTGAGGGAACTT	865
Db	1111		1111
QY	939	CTCCATGCGCCGAGCCAGACATGAGCGTGGCTTCCCTGTGGCACTGGAGCAGCTCAT	998
Db	1111		1111
QY	866	CTCCACGCGCAGCCAGACATGAGTGTGCTCTCTCTCCATGTGGAGCAGCTCAT	925
Db	1111		1111
QY	999	GATGGCCAGCCAGGAATGTGTGTTGAGGGCTCTCACCACTGCTCCATGGCCCC	1058
Db	1111		1111
QY	926	GATTGCCAGGCAAGGAATGCACTTCAAGGGCTTTTACTGCGAGCTCTGCCACACC	985
Db	1111		1111
QY	1059	CCAGACTGCTGCGCCAGCTGCGCTGGCGCAGAGGCGCCCGAGGTGGCAGCGGATA	1118
Db	1111		1111
QY	986	TGACATCTGTCCAGATCACTGACGTGGCCAGAGAGTGTCCAGGTGGCACTGAATA	1045
Db	1111		1111
QY	1119	CAGCTAGTGCACCGGACCATGGCCACGCCCTGTGGCTTCCCTGTGGCACTGGTCTCTG	1178
Db	1111		1111
QY	1046	TGGGCTTGTGACCGGGCCATGGCCAGCCACTGTCCGAGACTACCTGCTGCTGCTCTG	1105
Db	1111		1111
QY	1179	GACTGCCCTGTGATGTCAAGGCGAGTACTTCGGCTCCCTGGCCACTACCACTGAGC	1238
Db	1111		1111
QY	1106	GACTAACTGTGCGACGCTCAAGGTGTGACACTTCTGGCCCTGGCCCACTATATGACG	1165
Db	1111		1111
QY	1239	CATGGCTCTGCGACGGCTCCCGACGACCGAGGAGGCTCCCGACGACGACGAGGT	1298
Db	1111		1111
QY	1166	CATGGCTCTGCGAAGACCCACG---CCAAGGTGAATCTCGAAGCGAGGACAGT	1222
Db	1111		1111
QY	1299	CTTCTGACGCCCCCACTCTTAAGCCCGAGGCGCTGTGCTGCGCAGGAGCTGGA	1358
Db	1111		1111
QY	1223	CTTC---CAGCCCTCCACTCCCATGAGCCCTGGGCCCCACACTGCGCGACCCAGA	1279
Db	1111		1111
QY	1359	GGAGCGAGGAGCTTGGCAAGGCACACTGAGCGTGCCATCTCTGGGGCAGGAGGC	1418
Db	1111		1111
QY	1280	AGACCGAGGAGCTT-----GGAGGAGC	1304
Db	1111		1111
QY	1419	GCTGGGCTGCACGCGCTGTGCGGCTCTGCGCAGGTGGACCTGCTTCGGGCTGTGAT	1478
Db	1111		1111
QY	1305	GCTGGGCTTACACTTGTGCAGAGTCTGCGCAAGGTGGACCTGCTGCAGTGTGTGT	1364
Db	1111		1111
QY	1479	CTCCAGACGCTGCGAGCGCTCACTGGCAAGTATCGGAGCTGCAACCGTGAGGATGACTT	1538
Db	1111		1111
QY	1365	GACTCAGGCACTGCGCGCTCTCCCTGGCCAAAGTACTCACAGCTGCGCGCGAGGATGACTT	1424
Db	1111		1111
QY	1539	CTGTGAGCTGCGAGGCGCCGAGACATCCAGCTTAAGACCCACAGAGCCAGAGGCGCAG	1598
Db	1111		1111
QY	1425	CTTCAGGCTACGAGGCGCCAGACATCCACCCAGACTCACAGACCCAGAGGTGAG	1484
Db	1111		1111
QY	1599	GATGCCAGCTGTCTCCAGGGGAGGGCTGACATCTTCCATCGGCTGGGCGCCCTGTGTC	1658
Db	1111		1111
QY	1485	GATGCTAGCTGTGCGAGGTGAGGTGACTGACTTCTTCCATCGCTGGGACCTTGTGTC	1544
Db	1111		1111
QY	1659	TGTGTTCTCAGCCAGAACCGGTGGGCTGTGGGCGCCCTCCACTGACCCGAGGAGA	1718
Db	1111		1111
QY	1545	TGTGTTTTCAACCAAGAACCGATGCGAGTGTGTAGGGCGGTGCACTGACTCGAGGAGA	1604
Db	1111		1111

QY	1719	GGCGGCTTTGGCTTCAAGCTTTCGGGAGACTCGCTGTCTCATCGTGCCTCATTC	1778
Db	1111		1111
QY	1605	GGGAAGCTTTCGGGTTTCAAGCTTTCGGGAGACTTTCAGTCTTCTCATCGTCTGCTGCTCC	1664
Db	1111		1111
QY	1779	AGGAGCCAGGCGCGCGG-----	1795
Db	1111		1111
QY	1665	TGGGGCTCAGGCTGAGGTAAAGCCCTTAGCCCCGGGTGCTGGAATCTTAGAGGACCGC	1724
Db	1111		1111
QY	1796	-----CGGCTGGCTTGAAGGAGG	1814
Db	1111		1111
QY	1725	TAGATATGCCACCACTTCCAGATCCATTTCTTCGCACTCGCTGGCTTGAAGGAGG	1784
Db	1111		1111
QY	1815	CGACTACTTGTCTAGTGAATCGGAGCAGCATCAGGTGTGTGAGACACCGGAGAGTGT	1874
Db	1111		1111
QY	1785	CGACTACATGCTCTCGTGAACCGGAGCGCTGCAAGTGTGTGAAGCACCTGGAGGTGT	1844
Db	1111		1111
QY	1875	GACGAGCTGAAGCTTGGGAGAGGCGGCGCAGCTGCAAGTGTGTGCTGTGCTGCC	1934
Db	1111		1111
QY	1845	GACACAGCTGAGGAGCATGGCGAGGAGCGGTGAGCTCAGGTGTGTATTTGCTGCC	1904
Db	1111		1111
QY	1935	CAGCTCTAGACTTGCAGCTTGGGGGACCGCGCTCTGCTGTGGGCCCCAGGGGCT	1994
Db	1111		1111
QY	1905	CAGCTCTAGCAGCTGGCACAGGACCCCGAGCGGAGCACTGCTG-----	1950
Db	1111		1111
QY	1995	TCTAAGAGCAGAGGAGCATGTTTGCAGAGACCCCGGCACTCCACGTGGGCGCAGTCCCG	2054
Db	1111		1111
QY	1951	----TGAAACAGAGAGAGTGTGTTTGTAGACACCCGATGCCACACTACCGCCCTG	2006
Db	1111		1111
QY	2055	GCCCTCTCACTGGAGCGGAAGGCCAGGCGGCAAGACTGAGGCTGCCCCCAGAC	2114
Db	1111		1111
QY	2007	GCCATCTTGGCTGGAGCGCGAAGATAAACAAGGCAAGACTGGAAGCCACCTTGACCC	2066
Db	1111		1111
QY	2115	CTGTGCCCCAGTGAA 2129	
Db	2067	CTGACCAACAGGAA 2081	
QY	BD183400	Novel genes and proteins encoded by the genes.	
Db	BD183400	Novel genes and proteins encoded by the genes.	
QY	BD183400.1	GI:31875600	
Db	BD183400.1	GI:31875600	
QY	JP 2002345492-A/113		
Db	JP 2002345492-A/113		
QY	Homo sapiens (human)		
Db	Homo sapiens (human)		
QY	ORGANISM	Homo sapiens	
Db	ORGANISM	Homo sapiens	
QY	REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
Db	REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
QY	AUTHORS	Ohara, O., Nagase, T. and Nakajima, D.	
Db	AUTHORS	Ohara, O., Nagase, T. and Nakajima, D.	
QY	TITLE	Novel genes and proteins encoded by the genes	
Db	TITLE	Novel genes and proteins encoded by the genes	
QY	JOURNAL	Patent: JP 2002345492-A 113 03-DEC-2002;	
Db	JOURNAL	Patent: JP 2002345492-A 113 03-DEC-2002;	
QY	COMMENT	KAZUSA DNA RESEARCH INSTITUTE	
Db	COMMENT	KAZUSA DNA RESEARCH INSTITUTE	
QY	OS	Homo sapiens (human)	
Db	OS	Homo sapiens (human)	
QY	PN	JP 2002345492-A/113	
Db	PN	JP 2002345492-A/113	
QY	PD	03-DEC-2002	
Db	PD	03-DEC-2002	
QY	PF	26-FEB-2002 JP 2002049009	
Db	PF	26-FEB-2002 JP 2002049009	
QY	PI	OSAMU OHARA, TAKAHIRO NAGASE, DAISUKE NAKAJIMA	
Db	PI	OSAMU OHARA, TAKAHIRO NAGASE, DAISUKE NAKAJIMA	
QY	PC	C12N15/09, C07K14/47//A61K31/711, A61K38/00, A61K48/00, A61P25/00,	
Db	PC	C12N15/09, C07K14/47//A61K31/711, A61K38/00, A61K48/00, A61P25/00,	
QY	PC	A61P25/14, A61P35/00, C12N15/00, A61K37/02	
Db	PC	A61P25/14, A61P35/00, C12N15/00, A61K37/02	
QY	CC	Novel genes and proteins encoded by the genes FH	Key
Db	CC	Novel genes and proteins encoded by the genes FH	Key
QY	FT	Location/Qualifiers (518) . . (1540) .	
Db	FT	Location/Qualifiers (518) . . (1540) .	
QY	FEATURES	Location/Qualifiers	
Db	FEATURES	Location/Qualifiers	
QY	Source	1. 4729	
Db	Source	1. 4729	
QY	/organism="Homo sapiens"		
Db	/organism="Homo sapiens"		
QY	/mol_type="genomic DNA"		
Db	/mol_type="genomic DNA"		
QY	/db_xref="taxon:9606"		
Db	/db_xref="taxon:9606"		
QY	ORIGIN		
Db	ORIGIN		
QY	Query Match	36.8%; Score 798.8; DB 6; Length 4729;	
Db	Query Match	36.8%; Score 798.8; DB 6; Length 4729;	
QY	Best Local Similarity	92.1%; Pred. No. 3.7e-115;	
Db	Best Local Similarity	92.1%; Pred. No. 3.7e-115;	

Db	390	TGACTTCTCAGTCGTTCTCAAGGATTTTCATCTCGGAACATTACAGCGAGGACAGCTATCT	449
QY	579	CTACGAGGAGAAATCAGGAGCTGAGGCGCTTCGCGGCGAGGCCCATGCGGAGACCCCAAGCCG	638
Db	450	ATATGAAGACGAATAGCAGACCTGATGACCTGAGACAGGCTTGCGGACGCCACGCGC	509
QY	639	GAATGAGTCGGGCTGGAGTGTCTCAGAGCTATTACACAGCTGCTGCTTCCTGATGC	698
Db	510	GGATGGGCGGGTTGAACCTGCTCATGAGCTACTTTCATCCAGCTGGGATTTGTGGAGAG	569
QY	699	GCCTTCTCCCTCACCCCTGCCAGGAGCCTCGGGCTCTTCTTCCACTGGTACGACTCGCTTAC	758
Db	570	CCGCTTCTTCCCCCCCACCGACAGATGGGATCTTGTTTACCTGGTATGACTCCCTCAC	629
QY	759	TGGGCTCCCGCCAGCAGCGTGCCTTGGGCTTCGAGAGGGAGCGTCTCTTCAACAT	818
Db	630	CGGGTCCCGGTGAGCCAGCAGAATTTGCTGCTGGAGAGGCCAGCATTTCTGTTCACAT	689
QY	819	CGGTGCTCTCCACACGAGATTGGGGCGCGCCAGGACCGTCTCTGACCGAGGCTGCCCG	878
Db	690	CGGAGCCCTCTACACAGATCGGGACCCCGGTGCAATCGGGGACCCAACTGGGCTGGA	749
QY	879	CCGCGCTATGAGGCGCTTCCAGAGCGCGCTGGGGCTTCAGCCTCTCTGAGGAGAACTT	938
Db	750	CGGTGCGGTGAGCGCTTTCAGAGAGCGCGAGGGGTTTTCATCACCTGAAAGAGAGCTT	809
QY	939	CTCCATGCGCGGAGCCAGACATGAGGCTGCTGCTCTGCGCATGCGAGTGGAGCTCAT	998
Db	810	CACCTCACACTCCGAGTTACGACATGAGCCCTGCCATGCTCAGCGTGTGCTCAAGATGAT	869
QY	999	GATGGCCAGGCCACAGAAATGTGTTTGGAGGCTCTCTCACACCTGCTCCATGGCCCC	1058
Db	870	GCTTGCACAGCCAGAGAACGTTGTGAGAAATCTGCTCTCTG-----GGATCCG	923
QY	1059	CCAGACTGCTGCTGCCAGCTGCGCTGGCGCAGAGGCGCCCGAGGTGGAGCCGAGTA	1118
Db	924	GAATGAGTCTTCTGCTGCTGTTGAAGTGGCTCAGAGGCTGCCAAGGTGGGAGGTGTA	983
QY	1119	CAGCTAGTGACCGGACCATGGCGCAGCCAGCCCGTCCAGCACTACGTCGCTGCTCTCTG	1178
Db	984	TCGGCAGCTGCACACGGCCATGAGCCAGGCCCGGTGAAGAGAGAACATCCCTTACTCTG	1043
QY	1179	GACTGCTCTGTGATGTCAGGCGGAGTACTTCCGCTCCCTGCGCCACCTACCACTGATGC	1238
Db	1044	GGCCAGCTGCTGCTGCTGAGGCGCCACCATAGCGGCCCTGGCCACTACTTTCGAGC	1103
QY	1239	CATGGCCCTCTGCGACGCTCCCGACGACCGAGGAGAGTCCCGACGACGAGCGAGT	1298
Db	1104	CACCTTCTCATGACACCACTGTTGAAGCCTGGTGCATGAGGACCATCAGGAGAGTG	1163
QY	1299	CTTCTGAGC-----CCCCGACCTCTTAAGCCCGAGGCCCTGTGCTGCC--GCAGGA	1352
Db	1164	CTGTGCCAGCTCTACGACCACTGCCCCAGAGGAGCTGACGCTTGGCCACCCCTCAAGAG	1223
QY	1353	GCTGAGGAGCGCAGCGAGCTTGGCAAGCGACACTGTAAGCGTGCATCTCTGGGCGAGGA	1412
Db	1224	CGGCCACGAGCCGACAGCTGGGCAAGTCCCACTGCGAGAGCGGTGGCCCAACGGA	1283
QY	1413	GGAGCGCTGCGGCTGACGCGCTGTGCGGCTCTGCGGAGGTGGACCTGCTTCGGGC	1472
Db	1284	GGAGTCGTGAGGGAGGCGCAGCTGTGCAAGAGCTCCGCAACATCGAGGTGCTCCAGGA	1343
QY	1473	TGTGATCTCCAGACGCTGACGCTCTACTGCGCAAGTATGCGGAGCTGACCGGTGAGGA	1532
Db	1344	CGTGTGTCGCGGCGCAGAGCGGTCCCGCTCAAGTACGCGGACGACCGAGCAGGA	1403
QY	1533	TGACTTCTGTGAGGCTGCCGAGGCCCGGACATCCAGCTTAAGACCCACGAGGCCGAGA	1592
Db	1404	CGATCTGTCTCACTTGATGATGCTCCCGACATTTATCTTAAACTTGAGCAAGAGTTGA	1463
QY	1593	GGCCAGGATGCAACGCTCTCCAGGGAGGGGCTGACATCTTCCATCGCTGGGCGC	1652
Db	1464	AATTATATGCCACAGTTCTTCCAAAGGTGACAGCAACGGGACTTCTTCCAGAGCTGGGCGC	1523
QY	1653	CCTGCTCTGTTCTCAGCCAAAGAACCGGTGGCGGCTGGTGGGGCCCGTCCACCTGACCCG	1712
Db	1524	CCTGCTCAGTGTCTTTCGGCTTAACAGAGATGGAACACCTCCTCGAAGCATTCACCTTCACTGC	1583
QY	1713	AGGAGAGGGCGGCTTGGGCTTCACGCTTCGGGAGAGCTGCGCTGCTCTCATCTGCTGCCGT	1772
Db	1584	AGAAGAAGGGGATCTGGGGTTACCTTTGAGAGGAAACTTCCCACTGCAAGTCCATTTCTCT	1643
QY	1773	CATTCCAGGAGCCAGCGCGCGCGCTGSCCTGAAAGGAGGCGACTACATTTGTGTCAGT	1832
Db	1644	GGATCTCTCACTGCTCTGCTGCGCTGGCGGAGCCAAAGGAGGATTTATTTGTTCCAT	1703
QY	1833	GAATGGCGAGCCATGACAGGTGGTGGAGACACGCGGAGGTTGGTACGAGCTGAAGGCTGC	1892
Db	1704	TCAGACGTGGATTGCAAGTGGCTGACGCTGAGCGAGGTCATGAAGCTGCTGAAGGCGCTG	1763
QY	1893	GGAGAGGGCGGCGCGCAGCCTGACAGGTGGTGGTCTGCTGCTGCCAGCTC	1940
Db	1764	TGGCAGGGACGGCTGGAGATGAAGGTGCTGAGCCTCTTGGACTTCAC	1811
RESULT 11			
BC036447			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REMARK			
COMMENT			

Db 1299 GGAGTCGGTGGGAGCCAGCCTCTGCAAGAGCTGCGGAGCAATTGAGTGTCTACAGAA 1358
QY 1473 TGTGATCTCCAGACGCTGACGGCTCACTGGCCAAAGTATGCGAGCTGACCGTGAGGA 1532
Db 1359 GGTGCTGTGTCGCGACAGGAACGCTCCCGCTCACTGACGCCAGCACCAGGAGGAGGA 1418
QY 1533 TGACTTCTGTGAGCTGCGGAGGCGCCGGACATCCAGCCTAAGACCCAGAGCCAGA 1592
Db 1419 TGACCTGCTGAACCTGATCGACGCCCGCCAGTGTGTGCTTAAACTGAGCAAGAGTTGA 1478
QY 1593 GGCAGAGTCCAGCGCTGTCAGGAGGAGGCGCTGACATCTTCATCGCTGGGGCC 1652
Db 1479 CATTATATGCCCAGTCTCCAGCTGACAGTCAGTCAGGACTTCTTCAGAGCTGGGCC 1538
QY 1653 CTTGCTGTGTTCTCAGCCAAAGAACCGTGGCGGCTGGTGGGCGCCGTCACCTGACCCG 1712
Db 1539 CTTATCTGTGTTTTCGGCTAACAAGCGGTGGAAGCGCTCTCGAAGCATCGCTTCACTGC 1598
QY 1713 AGAGAGGGCGGCTTTGGCTCAGCTCGCTTCGGGAGACTCGCTGCTCTCATCGTGCCT 1772
Db 1599 AGAAGAGGGGACTTGGGGTTTCACTTTGAGAGGAAACGCCCGCTTCAGGTTCACTTCT 1658
QY 1773 CATTCCAGGAGGAGCGCGCGGCTGGCTGAAGGAGGCGGCTACATTTGTGTCAGT 1832
Db 1659 GATCTCTTACTGCTCTGCTCGGTGGCAGGAGCCCGGAGGAGATTATTTGCTCCAT 1718
QY 1833 GAATGGCGACGCTGAGGCTGGGAGACAGCGGAGGAGTGGTGGAGGCTGAGGCTGC 1892
Db 1719 TCAGCTGTGATTTGTAAGTGGCTGACGCTGAGTGAGGTTATGAAGCTGCTGAAGAGCTT 1778
QY 1893 GGGAGAGCGCGCGCGCTGCGAGTGGTGTGCTGCTGCTG 1932
Db 1779 TGGCGAGGACGAGATCGAGTGAAGTGTGAGGCTCTGCTGCTG 1818

RESULT 12
HSA347750
LOCUS HSA347750 2061 bp mRNA linear PRI 06-JAN-2003
DEFINITION Homo sapiens mRNA for p76RBE protein.
ACCESSION AJ347750
VERSION AJ347750.1 GI:23304861
KEYWORDS p76RBE protein.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Mircescu,H., Steuve,S., Savonet,V., Degraef,C., Mellor,H.,
Dumont,J.R., Maenhaut,C. and Pirson,I.
TITLE Identification and characterization of a novel activated RhoB
binding protein containing a PDZ domain whose expression is
specifically modulated in thyroid cells by cAMP
Eur. J. Biochem. 269 (24), 6241-6249 (2002)
JOURNAL 22361238
MEDLINE 12473120
PUBMED 2 (bases 1 to 2061)
REFERENCE
AUTHORS Pirson,I.
TITLE Direct Submission
Submitted (20-SEP-2001) Pirson I., Interdisciplinary Institute
IRIBHN, Université Libre de Bruxelles, 1070, Brussels, BELGIUM
JOURNAL
FEATURES
source 1..2061
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..2061
/function="putative RhoB effector"
/codon_start=1
/product="p76RBE protein"
/protein_id="CAC87939.1"
/db_xref="GI:23304862"
/db_xref="GOA:Q8IUC4"
/db_xref="SPTREMBL:Q8IUC4"

/translation="MTDALLPAAPQPLEKENDGYFRKCNPLAQTGRSKLQNRALIN
QOILKAVRMTGAENLLKVATNSKVRQVLELSFVNSDLQMLEKELEGISVGYQ
NTEAFIPLIPLIGLKETKDVFAVVLKDFILEHYSEDGYLVEDEIADLMDLRQART
PRDEAGVELLWTFYFQLQFVSRPEPTROMGLLATWVDSLTVGVPSQONLLLEKAS
VUFTALYQIGTRCDROTQAGLESAIDAQFQAGVLYNLKDTFTFTSYDMSAPML
SVLTKMLAQAGSVEFKISLPIRNEFFMLVKVQAQAAKGVYQOLHAAHQAPVK
ENIPYSWASLACVKAHYYAALHRYFTAILLIDHVPKPGTDLQHEKLSQLYDHMEPG
LTPLATLNQDQOROLKSHLRRAHHEESVREASLCKLRSIEYLOKVLCAQERS
RLTYAQHOBEDLLNLI DAPSVAKTEQVDIILPOFSKLTVDFFQKPLGSLVFSAN
KWTTPRSIRFTRAEEDLGTFLAGNAPVOVHFDLPYCSASVAGARSDYIVSLQVDC
KMLTJSEVKKLLKSGFDEIEMKVSVLLSTSSMHNKSATYSVGMQKTYSMICLAIDD
DKRTDKTKISKLSFLSWGNTKNROKASTLCLPSVGAARPQVKKLPSFFLLNSD
SSWY"

ORIGIN

Query Match 22.8%; Score 494.8; DB 9; Length 2061;
Best Local Similarity 56.7%; Pred. No. 1.2e-67;
Matches 976; Conservative 0; Mismatches 732; Indels 12; Gaps 3;
QY 219 GGGCTGTGACTCCCTGACGCGAGATCCAGTGCAGCGCAGCTGCAGAGCGCGAGGCCCCAGAT 278
Db 69 GGGCTGTATCCCTTGCACAAACCGCGGAGTAAATTGCAGAAATCAAAGAGCTGCTTT 128
QY 279 TCACCAGCAGATTGACAAGGAGCTGCAGATGCGGAGCGGCTGAGAACTCTACAGAGC 338
Db 129 GAATCAGCAGATCCTGAAAGCCGCTGCGGATGAGAGCGGAGCGGAAACCTTCTGAAAGT 188
QY 339 CACGAGCAACACCGGCTGAGAGAGCGCTGCGCTGGAGCTGAGCTACCTCAACTCCAA 398
Db 189 GGGCACAACCTCAAGGTGCGGAGCAAGTGCAGCTGAGCTGAGCTTCTGCAACTCAGA 248
QY 399 CCTGCAGCTGCTGAAGGAGGAGCTGAGGAGCTAGCGGTGGCGTGAACCTTGGCGCGCA 458
Db 249 CTTGCAGATGCTCAAGGAGAGCTGAGGAGGCTGAAACATCTCGGTGGCGCTTATCAGAA 308
QY 459 TGGGAGCGAAGCTGCTGCTGCCATGATCCCTCGGCTGAGGAGCTGAGGAGCCAGGAGCT 518
Db 309 CACAGAGGAGGCAATTAACGATTCCTCTTGGCGCTGAAGGAAACGAAAGAGCT 368
QY 519 GGACTGCTTACACCGCTGAAGGAGCTGATCTCAGTGCACCTTTTGAGAGGAGCGCGCTC 578
Db 369 CGACTTTGAGTCTGCTCTCAAGGATTTTATCTCGAATTTACAGTGAAGTGGCTATTT 428
QY 579 CTACGAGCGAGAAATCAGGAGCTGAGGAGCGCTCGGAGGCCATGCGAGCCCGAGCGG 638
Db 429 ATATGAAGATGAATTCAGATCTTATGATCTGAGCAAGCTTTGTGCGAGCGCTAGCGG 488
QY 639 GAATGAGTGGGCTGAGCTGCTGACAGCTTATCAACAGCTGCTGCTTCTCTGGATGC 698
Db 489 GGATGAGCGGGGTGGAATCTGATGACATATTCATCAGCTGGGCTTTGTGCGAG 548
QY 699 GCGCTTCTCACCCTCGCAGGAGCTCGGGCTCTTTCTTCCACTGGTACGACTCGCTTAC 758
Db 549 TCGATTTCTCCCGCCACACGCGAGATGGGACTCTGTTTACCTGTTATGATCTCTCAC 608
QY 759 TGGGGTCCCGGCCCGCAGCAGCGTGCCTGGCTTTCAGAAAGGCGAGCGTCTCTTCAACAT 818
Db 609 CGGGTTCCGGTTCAGCCAGCAGAGAACCTGCTGCTGAGAGAGCCAGTGTCTGTTCAACAC 668
QY 819 CGGTGCTCTCCACAGCAGATTTGGGCGCGCAGAGCCGCTCTCTCACCGAGGGTGC 878
Db 669 TGGGGCTCTTACACCCAGATTGGGACCCGGTGTGATCGGAGAGCGAGGCTGGGCTGGA 728
QY 879 CCGGCTATGGAGGCTTCCAGAGGCGCGCTGGGCTTTCAGCTCTTTCAGAGGAGAACTT 938
Db 729 GAGTGCATAGATGCTTTTCAGAGAGCGCGAGGGGTTTAAATTAACCTGAAAGACACAT 788
QY 939 CTCCCATGCGCGAGCCCGCAGACATGAGCGCTGCGCTCTCTCGGACCTGGAGAGCTCAT 998
Db 789 TACCCATATCTCAAGTTACGACATGAGCCCTGCCATGCTGAGCGTCTGCTGCTCAAAATGAT 848
QY 999 GATGCGCCAGCGCCAGGATGTGCTTTGAGGGCTCTCACACCTGCTCTCATGCGCCC 1058

```
Db      849  GCTTGACAAAGCCCAAGAAAGCGTGTGAGAAATCAGCCTTCCTG-----GGATCCG 902
QY      1059  CCAAGACTGCTGGCCAGCTGGCCCTGGCGGAGAGGCGCCAGCGAGTGGCAGCCGAGTA 1118
Db      903  GAATGAATTTTCATGCTGGTGAAGGTGGCTCAGAGGCTGCTAAAGTGGGAGAGGTCTA 962
QY      1119  CAGGCTAGTGCACGGACCATGCCAGCCACCGCTCCAGCACTACTGCTGCTCTCTCTG 1178
Db      963  CCAACAGCTACAGCGACCATGAGCAGGCGCGGTGAAAGAAATCCCTTACTCTCTG 1022
QY      1179  GACTGCCCTTGTGTCATGTCAAGCCCGAGTACTTCCGCTCCCTGGCCCACTTACCAGTAGC 1238
Db      1023  GGCAGCTTAGCTGCTGAAGGCCACCACTACGCGGCCCTGGCCCACTACTTCACTGC 1082
QY      1239  CATGGCCCTCTGGACGGCTCCCGACGACCGAGGAGAGCTCCCAACGACGAGCAGGT 1298
Db      1083  CATCTCTCTCATGACCAACAGGTGAAGCCAGGACCGGATCTGGACCAACGAGAGAAGTG 1142
QY      1299  CTTCTCTGCAGC---CCCCACCTCTCTAAGCCCGAGGCCCTGTGC--TGCAGCAGGA 1352
Db      1143  CTTGTCCTCAGCTCTAGACCAATGCGAGGGGCTGACACCTTGGCCCACTGAAGAA 1202
QY      1353  GCTGAGGAGCGCAGCAGCTTGGCAAGGCACACTGAAGCGTGCCTCTCTGGCGCAGGA 1412
Db      1203  TGATCAGCAGCGCGACAGCTGGGGAAGTCCCACTTGGCAGAGCCATGGCTCATCAGA 1262
QY      1413  GGAGGCGCTGGCGCTGCAGCCCTGTGCGCGGTCTTGGCGAGGTGGACCTGCTTCGGCG 1472
Db      1263  GGAGTGGGTGGCGGAGGCGAGCTCTGCAAGAAGCTGCGGAGCATTGAGGTGCTACAGAA 1322
QY      1473  TGTGATCTCCAGAGCGCTCAGCGCTCACTGGCCAGTAGTGGAGCTCGACCGTGAGGA 1532
Db      1323  GTGTGCTGTGCGGCACAGGAAGCTTCCCGGTTCAGTTCAGCCCGACACCGAGGAGGA 1382
QY      1533  TGACTTCTGTGAGGTGCGAGGCGCCGACATCCAGCCTAAGACCCACCAAGCCAGA 1592
Db      1383  TGACTGCTGAACCTGATCGACGCCCCCAGTGTGTGCTTAAACTGAGCAAGAGTTGA 1442
QY      1593  GGCCAGGATGCCAGCTGTCCAGGGGAAGGGGCTTGACATCTTCCATPGGCTGGGCGC 1652
Db      1443  CATTAATATGCCCCAGTTCTCCAAGCTGACATCAGGACTTCTTCCAGAACTGGGCGC 1502
QY      1653  CTTGCTGTGTCTCAGCAAGAACCGGTGGCGGTGGTGGGCGCGTCCCTCATCGCTGCCGT 1712
Db      1503  CTTATCTGTGTTTCGGCTTAACAGCGGTGGAGCGCTCTCTCAAGCATCCGCTTCACTGC 1562
QY      1713  AGGAGAGGCGGCTTTGGCTCAGCCTTCGGGAGACTCGCCTGTCTCATCGCTGCCGT 1772
Db      1563  AGAAGAGGGGACTTGGGGTTTCACTTGAGAGGGAACGCCCGCTTCAGGTTCACTTCT 1622
QY      1773  CATTCAGGGAGCCAGCGCGGCTGGCTGGCTGAAGGAGGCGACTACATTTGTGTCAGT 1832
Db      1623  GGATCTTACTGCTGCTGCTCGGTGGAGGAGCCCGGGAAGAGATATATTTGTCTCCAT 1682
QY      1833  GAATGGGCGAGCATCAGCTGTGGAGACACGCGAGGTGGTGGAGGAGCTGAAGGCTGC 1892
Db      1683  TCAGCTTGTGATTGTAGTGGCTGACCTGAGTGAGGTATGAAGCTGCTGAAGAGCTT 1742
QY      1893  GGGAGAGCGGCGCCAGCCTGCAGGTGTGTGCTGCTGCTG 1932
Db      1743  TGGGAGGACGAGATCGAGATGAAGTGTGAGCCTCTG 1782
```

```
RESULT 13
AX835273 LOCUS 3500 bp DNA linear PAT 15-DEC-2003
DEFINITION Sequence 2397 from Patent EP1347046.
ACCESSION AX835273
VERSION AX835273.1 GI:39921408
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Isogai, T., Sugiyama, T., Otsuki, T., Wakamatsu, A., Sato, H., Ishii, S., Yamamoto, J. I., Isono, Y., Hio, Y., Otsuka, K., Nagai, K., Irie, R., Tamechika, I., Seki, N., Yoshikawa, T., Otsuka, M., Nagahari, K. and Masuno, Y.
TITLE Full-length cDNA sequences
JOURNAL Patent: EP 1347046-A 2397 24-SEP-2003; Research Association for Biotechnology (JP)
FEATURES
source location/Qualifiers
1..3500 /organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 22.8%; Score 494.8; DB 6; Length 3500;
Best Local Similarity 56.7%; Pred. No. 1e-67;
Matches 976; Conservative 0; Mismatches 732; Indels 12; Gaps 3;
QY 219 GGGCTGTGACTCCCTGAGCGAGATCCAGTGGCGGAGCTGCAGAGCGCGAGGCCAGAT 278
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 105 GGGCTGTATCCCTTGCACAAACCGGCCGGAGTAAATTCAGAAATCAAGAGCTCTTT 164
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 279 TCACCAGCAGATTGACAGGAGCTGCAGATGCGGAGCGGCTGAGAACTCTACAGAGC 338
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 165 GAATCAGCAGATCTCGAAGCCATCGGATGAGGACCGGAGCGGAAACCTTCTGAAAT 224
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 339 CACCAGCAACACGGGTGAGAGAGCGGTGCGCCCTGGAGCTGAGCTCAACTCCAA 398
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 225 GGCACAAACTCAAAGTGGCGGAGCAAGTGGGCTGGAGCTGAGCTTCTGCTCAACTCAG 284
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 399 CTGCACTGCTGGAAGGAGGAGCTGAGAGAGCTCAGCGTGGCGTGGACCTTGGCGGCA 458
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 285 CTTGAGATGCTCAAGAGAGAGCTGGAGGGCTGAACATCTCGGTGGGCTCTATCAGAA 344
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 459 TGGGAGCGAAGCTGTCACTGTCCCATGATCCCTTGGGCTTGAAGGAGACCAAGAGCT 518
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 345 CACAGAGGAGGATTTACGATTTCCCTGATTTCTTGGCTTGAAGAAACGAAGAGCT 404
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 519 GGAAGTGTCTACACCGCTGAGGAGCTGATCTCAGTGCACTTTGGAGAGGACGGCGCTC 578
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 405 CGACTTTGAGTGTCTCAAGGATTTTATCTCGGAAATTTACAGTGAAGATGGCTATTT 464
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 579 CTACGAGCAGAAATCAGGAGCTGAGGCGCTGCGGAGGAGCCATCGGACCCCGAGCG 638
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 465 ATATGAAGATGAATGAGATCTTATGATCTGAGACAAGCTTGTGAGAGCGCTTAGCGG 524
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 639 GAATGAGTGGGCTTGGAGCTGCTCACAGCCTATTACAAACAGCTGTGCTTCTCGATGC 698
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 525 GGATGAGCGCGGGTGGAACTGCTGATGACATCTTCACTCAGCTGGGCTTGTGAGAG 584
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 699 GGGCTTCTCACTCCCTGCGAGAGCTCGGCTCTTCTTCCACTGGTACGACTGCTTAC 758
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 585 TCGATTTCTTCCGCGCCACACGCGAGATGGGACTCTCTTCACTGGTATGACTCTCTC 644
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 759 TGGGCTCCCGCCCGAGAGCGTGGCCCTTCGAGAGGAGCGGCTTCTTCAACAT 818
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 645 CGGGTTCGGTTCAGCCAGAGAACCTGCTGCTGGAGAGGCCAGTGTCTCTTCAACAC 704
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 819 CGGTGCCCTTCCACACGAGATTTGGGCGCGCCAGAGCGCTCTCTGCAACGAGGTCGCC 878
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 705 TGGGCGCTCTACACCCAGATTGGACCCCGTGTGATCGGAGAGCGGCTGGGCTGGA 764
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 879 CCGGCTATGAGAGCTTTCAGAGGCGCTTGGGCGCTTACGCTCTTGAGGAGAACTT 938
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 765 GAGTGCATAGATGCTTTCAGAGAGCGCGAGGGGTTTAAATTTACCTGAAAGACACATT 824
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 939 CTCCTATGCGCGGAGCCAGACATGAGCGTGTCTTCCCTCTGCGCACTGGAGAGCTCAT 998
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 825 TACCCATCTCCAAAGTTACGACATGAGCCCTGCGCATGCTCAGCGTGTCTGCTCAAAATGAT 884
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 999 GATGGCCCGCCCGAGGAATGTGTGTTGAGGGCTCTCTACCACTGCTCTCATGGCCCC 1058
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Db 885 GCTGACAGCCAGCAAGAGAGGCTGTTGAGAAATCAGCCTTCCTG-----GGATCCG 938

QY 1059 CCAAGACTGCTGGCCAGCTGCGCTGGCGCAGGAGCGCCGACAGTGGCAGCCGAGTA 1118

Db 939 GAATGAATCTTCATGCTGCTGAAGTGGCTCAGGAGGCTGCTAAGTGGAGAGTCTA 998

QY 1119 CAGGCTAGTGCACCGGACCAATGCGCCAGCCAGCCAGCTCAGCAGTACGCTGCTCTCTG 1178

Db 999 CCAACAGCTACACGACGACATGAGCCAGGCGCGGTGAAGAGAAATCACCCTACTCCTG 1058

QY 1179 GACTGCGCTGCTGATGCTCAAGCGGAGTACTTCGCTCCGCTGCGCCACTACACAGTACG 1238

Db 1059 GGCAGCTTAGCCTGCGTGAAGGCCACCACTACGCGGCCCTGCCCCACTCTTCACTGC 1118

QY 1239 CATGGCCCTCTGGACCGCTCCCGACGCGACCGGAGAGCTCCCGACGACAGCAGAGGT 1298

Db 1119 CATCTCTCTCATCGACCAACAGGTGAAGCCAGGACGCGATCTGACCAACAGGAGAAGTG 1178

QY 1299 CTTCTCTGACG-----CCCCACCTCTCTTAAGCCCGGAGGCGCTGTGC--TGCGGACGA 1352

Db 1179 CTTCTCTGACGCTTACGACCACTACGACGAGGCGCTGACACCCCTTGCCACACTGAAGA 1238

QY 1353 GCTGGAGGCGGAGGAGCTTGGCAAGGCACACTGAAGCGTGCCTATCTGGGGCAGGA 1412

Db 1239 TGATCAGCAGCGCGACAGCTGGGGAAGTCCCACTTGGCAGAGCCATGGCTCATCAGA 1298

QY 1413 GGAGCGCTGCGGTGACGCGCTCTGCGCGCTCTGCGGAGGTGAAGCTGCTTCGGCG 1472

Db 1299 GGAGTGGTGGGAGGCGAGCCTCTGCAAGAGCTGCGGAGCTTGAAGTGTCTACAGAA 1358

QY 1473 TGTGATCTCCAGACGCTGACGCTCTACTGGCCAGATATGCGAGCTCGACCTGAGGA 1532

Db 1359 GGTCTGTGTGCGCCACAGGAACCTCCCGCTCACGTACGCTGCGCAGCACAGGAGGAGGA 1418

QY 1533 TGACTCTGTGAGGCTGCGGAGGCGCGGACATCAGCCTTAAGACCCACAGAGCCAGA 1592

Db 1419 TGACCTGCTGAACTGATCGAGCGCCCGAGTGTGTGCTTAAGCTGAGCAAGAGTTGA 1478

QY 1593 GGCAGAGATGACGAGCTGCTCCAGGGAAGGGCGCTGACATCTTCCATCGGCTGGGGCC 1652

Db 1479 CATATATTTGCCCCAGTTCTCAAGCTGACAGTCAAGCTTCTTCCAGAGCTGGGCGCC 1538

QY 1653 CTTGCTGTGTTCTCAGCAAGAACCGTGGCGGCTGTTGGGCGCCCGTCCACCTGACCG 1712

Db 1539 CTTATCTGTGTTTTCGGCTAACAGCGGTGAGCGCTCTCTGAAAGCATCCGCTTCACTGC 1598

QY 1713 AGGAGAGGCGGCTTTGGCTTCAAGCTTGGGAGAGCTGCGCTGTCTCTCATCGCTGCGGT 1772

Db 1599 AGAAGAGGAGACTTGGGGTTCACTTGAAGAGGNAAGCCCGCTTCAAGTTCACTTCT 1658

QY 1773 CATTCAGGAGCCAGCGCGGCTGGCTGAGAGGCGGAGCTACATTTGTGCTAGT 1832

Db 1659 GGAATCTTACTGCTCTGCTCGTGGCAGGAGCGCGGAGAGATATATTTGTCTCCAT 1718

QY 1833 GAATGGGAGCCATGACAGTGTGTGAGACACGCGAGGTGTGAGGAGCTGAGGCTGC 1892

Db 1719 TCAGCTTGTGATTGAAGTGGCTGACGCTGAGTGAAGTTATGAAGCTGCTGAAGAGCTT 1778

QY 1893 GGGAGAGCGGCGCGCAGCTGACGCTGAGTGTGTGCTGCTG 1932

Db 1779 TGGGAGGACGAGATCGAGATGAAGTGTGAGCTTCTG 1818

RESULT 14
AK098246
LOCUS AK098246 3500 bp mRNA linear PRI 15-JUL-2002
DEFINITION Homo sapiens cDNA FLJ40927 fis, clone UTERU2006593, weakly similar to GTP-RHO BINDING PROTEIN 1.
ACCESSION AK098246
VERSION AK098246.1 GI:21758222
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS Oshima,A., Takahashi-Fujii,A., Tanase,T., Imose,N., Takeuchi,K., Arita,M., Musashino,K., Yuuki,H., Harai,H., Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K., Wagatsuma,M., Murakawa,K., Kanehori,K., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S., Nagahari,K., Masuno,Y., Nagai,K. and Isogai,T.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
2 (bases 1 to 3500)
AUTHORS Isogai,T. and Yamamoto,J.
TITLE Direct Submission
JOURNAL Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 252-0812, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB.
FEATURES
source Location/Qualifiers
1. 3500
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UTERU2006593"
/tissue_type="uterus"
/clone_lib="UTERU2"
/note="cloning vector: pME18SFL3"

ORIGIN
Query Match 22.8%; Score 494.8; DB 9; Length 3500;
Best Local Similarity 56.7%; Pred. No. le-67;
Matches 976; Conservative 0; Mismatches 732; Indels 12; Gaps 3;

QY 219 GGGCTGTGACTCCCTGACGAGATCCAGTGGCGGAGCTGAGAGCCGAGGCGGAGAT 278

Db 105 GGGCTGTGATTCCTTGCACAAACCGCGCGAGTAAATTGAGAATCAAGAGCTGCTT 164

QY 279 TCACAGCAGATTGACAGAGCTGAGATCGGACCGGCGCTGAGACCTTACAGAGC 338

Db 165 GAATCAGCAGATCTGAAAGCCATCGGATGAGGACCGGAGGAAACCTTCTGAAAGT 224

QY 339 CACAGCAACAACCGGTGAGAGAGCGTCCCTGGAGCTGAGCTACGTCACTCAA 398

Db 225 GGCCACAACACTCAAAAGTGGCGGAGCAAGTGGCGCTGAGCTGAGCTCGTCACTCAGA 284

QY 399 CCTGCAGCTGTGAAGGAGGAGCTGAGGAGCTCAGCGGTGGCGTGGAGCCCTGGCGCGCA 458

Db 285 CCTGCAGATCTCAAGGAGAGCTGGAGGGCTGAACATCTCGGTGGCGCTTATCAGAA 344

QY 459 TGGGAGCGAAGCTGTCACTGTCCCATGATCCCTGGCGCTGAGGAGACCAAGAGCT 518

Db 345 CACAGAGGAGGCAATTTACGATTCCCTGATTCCTCTTGGCTTGAAGGAAACGAAGAAGT 404

QY 519 GGACTGTGCTTACACCGCTGAAGGAGCTGATCTCAGTGCATTTTGGAGAGGACGGCGCTC 578

Db 405 CGACTTTCAGCTCGTCTCTCAGGATTTTATCTGGAAATTTACAGTGAAGATGGCTATT 464

QY 579 CTACGAGCAGAAATCAGGAGCTGAGGAGCTGGAGCCCTGGCGGAGGCCATCGGACCCCGCG 638

Db 465 ATATGAAGATGAATTTGAGATCTTATGGATCTGAGACAAGCTTGTTCGAGACCTTAGCG 524

QY 639 GAATGAGTGGCGCTGGAGCTGCACAGCCCTATTACAAACAGCTGTGCTTCTCGATGC 698

Db 525 GGATAGGCGCGGGTGGAACTGCTGATGACATATTCATCCAGCTGGGCTTTGTGAGAG 584
Qy 699 GCGCTTCTCACCCTGCGAGAGCCCTCGGGCTCTTCTTCCACTGGTACGACTCGCTTAC 758
Db 585 TCGAATCTTCCGCGCCACACGCGAGATGGGACTCCTGTTCACCTGGTATGACTCTCTAC 644
Qy 759 TGGGGTCCCGGCGCCAGCAGCGTGCCTCGGCTTCGAGAAAGGCGAGCTTCTTCAACAT 818
Db 645 CGGGGTTCGGGTGAGCAGCAGAACTGCTGCTGGAGAGGCGCAGTGTCTGTTCACAC 704
Qy 819 CGGTGCCCTCCACACGAGATTTGGGGCGCGCCAGACCGCTCCTGCACCGAGGCTGCCG 878
Db 705 TGGGGCCCTCTACACCCAGATTGGGACCCCGGTGTGATCGGACGACGAGCTGGGCTGGA 764
Qy 879 CCGGCTATGGAGCCCTTCAGAGGGCGCTGGGGCCCTTCAGCCTCTCGAGGGAGAACTT 938
Db 765 GAGTGCATAGATAGCTTTTCAGAGAGCGCGAGGGTTTAAATTAACCTGAAGACACATT 824
Qy 939 CTCCCATGCGCGGAGCCAGACATGAGCGTGGCTCCCTCTCGGCACTGGAGCAGTCA 998
Db 825 TACCCATCTCCAAATTACGACATGAGCCCTGCCATGCTCAGCGTGTCTGTCAAAATGAT 884
Qy 999 GATGGCCCGAGCCCGAGAAATGTGTGTTGAGGGCTCTCACACCTGCGCTCCATGGCCCC 1058
Db 885 GCTTGACAAAGCCAAAGAAAGCTGTTTGAGAAATCAGCCTTCCTG-----GATCCG 938
Qy 1059 CAGAGTGGCTGGCCAGTGGCCCTGGCGAGGAGCGCCGACAGTGGCAGCCGAGTA 1118
Db 939 GAATGAATCTTTCATGCTGGTGAAGTGGCTCAGAGAGGCTGTAAAGTGGGAGAGTCTA 998
Qy 1119 CAGGCTAGTGCACCGACCATGSCCAGACCCAGCCGTCACGACTAGTGGCTGTCTCTG 1178
Db 999 CCAACAGCTACACGACCATGAGCCAGGCGCGCGTGAAGAGAAATCCCTACTCTCTG 1058
Qy 1179 GATGCCCTGGTGCATGTCAAGGCCGAGTACTTCGCTCCCTGGCCCACTACCACTAGC 1238
Db 1059 GGCAGCTTAGCTGGTGAAGSCCACCACCTAGCGGCGCTGGCCCACTACTTCACTGC 1118
Qy 1239 CATGGCCCTTGGAGCGCTCCGACGACCGAGGAGAGTCCCCACGACGAGCAGT 1298
Db 1119 CATCTCTCTATCGACACCACTGAGGTGAAGCGAGGCAAGATCTGGACCACTGAGAGA 1178
Qy 1299 CTTCTCTGACG-----CCCCACCTCTCTTAAGCCCGAGGCGCTTGGC--TGCGCAGGA 1352
Db 1179 CTGTCTCCAGCTCTAGACCACTGCGAGGGGCTGACACCCCTTGGCCACACTGAAGA 1238
Qy 1353 GCTGAGGAGCGCAGGCTGTTGGCAAGGCAACCTGAAAGCGTGCATCTCTGGGCGAGA 1412
Db 1239 TGATCAGCAGCGCCGACAGCTGGGAAGTCCCACTTGGCAGAGCCATGGCTCATCAGA 1298
Qy 1413 GGAGGGCTGCGGCTGCAGCCCTGTGCGGCTCCTGCGGAGGTGGACCTCTTGGGCG 1472
Db 1299 GGAGTCGGTGGGAGGCGAGCTCTGCAAGAAAGCTTGGCGAGCATTTGAGGTCTACAGA 1358
Qy 1473 TGTGATCTCCACAGCGCTGACGCTCACTGCGCAAGTATGCGGAGCTGCACCGTGAAGA 1532
Db 1359 GGTGCTGTGCGGCACAGGAACGTTCCGGCTCAGTACGCCCCAGACACAGAGGAGGA 1418
Qy 1533 TGACTTCTGTGAGGTGCGGAGGCCCGGACATCCAGCCTTAAGACCCACAGAGCCAGA 1592
Db 1419 TGACCTGTGAACCTGATCGAGCGCCCGCAGTGTGTGTGTTAAACTGAGCAGAGATTGA 1478
Qy 1593 GGCCAGGATGCCACGCTGTCCAGGGAGAGGGGCTTGACATCTTTCATTCGGCTGGGGCC 1652
Db 1479 CATTAATATGGCCCAAGTTCTCCAAGCTGACATCAAGGACTTCTTCCAGAAAGCTGGGCC 1538
Qy 1653 CTTGCTGTGTCTCAGCAAGAAACCGGTGGCGGTGGTGGGGCCCGTCCACCTGACCCG 1712
Db 1539 CTTATCTGTGTTTCGGCTAACACAGCGGTGAGCGCTCTCTGAAGCATCCGCTTCACTGC 1598
Qy 1713 AGGAGGCGGCTTTGGCTTACCGCTTCACGCTTCGGGAGACTCGCTGTCTTCATCGCTGCGGT 1772
Db 1599 AGAAGAGGGACTTTGGGTTTACCTTTCAGAGGGAACGCCCGCTTCAAGTTCACTTCT 1658

Qy 1773 CAUATCCAGGAGCAGCGCGCGGCTGCTGAAAGGGCGGCACTACATTTGTCTCAGT 1832
Db 1659 GGAATCCTTACTGTCTCTGCTCGGTGGCAGGAGCCCGGGAAGGAGATTATATTGTCTCAT 1718
Qy 1833 GAATGGGACCCATGCAAGTGTGGTGGAGACACGCGGAGTGGTACCGAGCTGAAGGCTGC 1892
Db 1719 TCAGCTTGTGGATTGTAAGTGGCTGACGCTGAGTGGATTATGAAGCTGTGAAGACTT 1778
Qy 1893 GCGAGAGGCGGGCGCCAGCCTGCAAGTGGTGTCTGCTGCTG 1932
Db 1779 TGGCAGGAGCAGATCGATGAAAGTCTGTGAGCCTCCTG 1818

RESULT 15

AX077672
LOCUS AX077672
DEFINITION Sequence 127 from Patent WO0105970.
ACCESSION AX077672
VERSION AX077672.1 GI:13122054
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Yue H., Tang, Y.T., Bandman, O., Hillman, J.L., Lal, P., Au-Young, J.,
Reddy, R., Yang, J., Baughn, M.R., Lu, D.A., Azimzai, Y. and
Patterson, C.

REFERENCE
AUTHORS
TITLE Gtp-binding protein associated factors
JOURNAL Patent: WO 0105970-A 127 25-JAN-2001;
Incyte Genomics, Inc. (US)

FEATURES
Location/Qualifiers
source
1..3019
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 3357264CB1"

ORIGIN
Query Match 22.5%; Score 487.8; DB 6; Length 3019;
Best Local Similarity 56.9%; Pred. No. 1.3e-66;
Matches 957; Conservative 0; Mismatches 713; Indels 12; Gaps 3;

Qy 257 TGCAGAGCGGAGGCCAGATTCACAGCAGATTCAGAGAGCTGCAAGAGCTGCAGATGCGGACGG 316
Db 190 TGCAGATCAAGAGCTGCTTTGAATCAGCATCTCTGAAAGCCGTGGCGATGAGGACCG 249

Qy 317 GCGCTGAGAACCTCTACAGAGCCACAGCAACACCGGTGAGAGACGGTCGCCCTGG 376
Db 250 GAGCGAAACCTCTGAAAGTGGCCACAACTCAAGGTGCGGAGCAAGTGGCGCTGG 309

Qy 377 AGCTGAGTACGTCAACTCCAACTGCGAGCTGTGAGAGAGAGCTGAGGAGCTCAAGC 436
Db 310 AGCTGAGTTCGTCAACTCAGACCTGCGATGCTCAAGGAAAGAGCTGAGGAGCTGAACA 369

Qy 437 GTGGGTGACCCCTGGCGGCGATGGGAGCGAGCTGCTCACTGTCCTCCATGATCCCTCTGG 496
Db 370 TCTCGTGGGCTGTATCAGAACACAGAGGAGCATTTACGATTCCTCTGATTCCTCTTG 429

Qy 497 GCCTGAAGGAGACCAAGAGCTGGAGCTGTCTACACCCCTGAAGAGCTGATCTCAGTGC 556
Db 430 GCCTGAAGGAAACGAAAGACGCTGCAGCTTTGCACTGCTCTCAAGGATTTTATCTGGAA 489

Qy 557 ACTTTGAGAGAGCGGCCCTCTACAGGAGCAAAATCAGGAGCTGGAGCCCTTGGCGC 616
Db 490 ATTACAGTGAAGATGGCTATTATATGAAGTGAATTCAGATCTTATGATCTTGAGAC 549

Qy 617 AGGCCATGCGGACCCCGAGCGGAGTGAAGTGGCGGCTGGAGCTGTCAAGCCTATTACA 676
Db 550 AAGCTTGTGCGACGCCCTAGCCGGATGAGCGCGGGTGGAACTGTGATGACATCTTCA 609

Qy 677 ACCAGCTGTGCTTCTCGATGCGGCTTCTCTCACCCCTGCCAGGAGCCTCGGGCTCTTCT 736

1684 CCCCCCTTCAGGTTCACTTCTCTGATCCTTACTGTCTCTGCTCGGTGGCAGGAGCCCGGG 1743
1811 AGGGCCACTACATTTGTTGTCAGTGAATGGGAGCCATGCGAGTGTGTCAGACACGCGGAGG 1870
1744 AAGGAGATTATATGTCTCCATTTCAGCTTGTGATTGTAAGTGGCTGACGCTGAGTGAGG 1803
1871 TGGTGAACGAGCTGAAGGCTTGGCGGAGAGCGCGCGCCAGCCTTGCAGGTTGGTGTGCTGC 1930
1804 TTATGAAGCTGCTGAAGAGCTTTTGGCGAGGACGAGATCGAGATGAAAGTCTGTGAGCCTCC 1863
1931 TG 1932
1864 TG 1865

Search completed: July 13, 2004, 06:10:05
Job time : 5576.17 secs

610 TCCAGCTGGGCTTTGTCGAGAGTCGATTTCTCCGGCCACACGCGAGATGGGACTCCTGT 669
737 TCCACTGGTACGACTCGCTTACTGGGTCGCCGCCAGCAGCGCTGCGCTGCCCTCGAGA 796
670 TCACCTGGTATGACTCTCTCACCGGGTTCGGGTACGCCAGCAGAACCTGCTGCGAGA 729
797 AGGGCAGCGTCTCTTCAACATCGGTGCCCTCCACAGCGAGATTGGGGCGCGCCAGGACC 856
730 AGGCCAGTGTCTGTTCAACACTGGGGCCCTCTACACCCAGATTGGGACCCGGTGTGATC 789
857 GCTCCTGCACCGAGGGTCCCGCGCTATGAGGCTTCCAGAGGGCGCTGGGGCCT 916
790 GGCAGACGCGAGGCTGGGCTGGAGAGTGCCTATGATGCTTTTCAGAGAGCCGCGAGGGTTT 849
917 TCAGCCTTCTGAGGGAGAACTTCTCCCATGCGCGAGCCCGAGCATAGAGCGCTGCGTCCC 976
850 TAAATTACCTGAAAGACACATTTACCCATCTCCAAAGTTACGACATGAGCGCTGCCATGC 909
977 TGTGCGCAGCTGGAGCAGCTCATGATGGCCCGAGGCCAGGAATGTGTGTTTGGGGCCTCT 1036
910 TCAGCGTGTCTCGTCAAAATGATGCTTGCACAAGCCCAAGAAAGCGTGTTTGAGAAATCA 969
1037 CACACCTGCTCCATGCGCCGCCCAAGACTGCTGCGCCAGCTGCGCTGGCGCAGGAGG 1096
970 GCCTTCTG-----GGATCNGAATGAATTCATCTGCTGGTGAAGGTGGCTCAGGAGG 1023
1097 CCGCCAGGTGGCAGCGAGTACAGGCTAGTGCACCGGACCATGGCCAGCAGCCCGTCC 1156
1024 CTGCTAAGGTGGAGAGTCTTACCAACAGCTACACGAGCCTAGCGAGCCTAGCGCGCGGTGA 1083
1157 ACGACTAGTGCCTGTCTCTGGACTGCCCTGGTGCATGTCAAGGCCGAGTACTTCCGCT 1216
1084 AAGAGAACATCCCCTACTCTCGGGCCAGCTTAGCTGCTGCTGAGGGCCACCACTACGGGG 1143
1217 CCCTGGCCCACTACACGTAGCCATGGCCCTCTGCGACGGCTCCCGCAGCGACCGAGGGAG 1276
1144 CCCTGGCCCACTACTTCTACTGCCATCTCTCTCATCGACCACAGGTGAAGCCAGGCAGG 1203
1277 AGCTCCCCACGACGAGCAGGTCTTCTGTCAGC-----CCCCCACTCTCTTAAGCCCCGA 1332
1204 ATCTGGACCAACGAGAGAGTGCCTGTCCAGCTCTACGACCAATGCCAGAGGGGTGA 1263
1333 GGCCTGTGC--TGCCGAGAGCTGGAGAGCGCGAGCTTGGCAAGGCACACCTGA 1390
1264 CACCTTGGCCACACTGAAGAATGATCAGCAGCGCGGACAGCTGGGGAAGTCCCACTTGC 1323
1391 AGCGTGCCATCTGGGGCAGGAGGCGCTGCGGCTGCACGCCCTGTGCGCGCTCTGC 1450
1324 GCAGAGCCATGGCTCATCACGAGAGTGGTGGGAGCGAGCGCTCTGCAAGAAGCTGC 1383
1451 GCGAGGTGGAAGCTGTGGGCTGTGATCTCCAGACGCTGACGCGCTCACTGGGCCAAGT 1510
1384 GGACGATTGAGGTGCTACAGAAGTGTCTGTGCGCCACAGGAACGCTCCCGGCTCACGT 1443
1511 ATGCGGAGCTGCACGCTGAGGATGACTTCTGTGAGGCTGCGAGGCCCGGACATCCAGC 1570
1444 AGCCCGACGACGAGGAGGATGACCTGCTGAACCTGATCGACGCCCCAGTGTGTTG 1503
1571 CTAAGACCCACACAGAAGCCAGAGGCCAGGATGCCACGCTGTCCAGGGGAAGGGCCCTG 1630
1504 CTAAACTGAGCAAGAGGTTGACATTATATGCCCCAGTTCTCCAAAGCTGACAGTCACGG 1563
1631 ACATCTTCATCGGCTGGGGCCCTGTCTGTGTTCTCAGCAGAACCGGTGGCGGCTGG 1690
1564 ACTTCTTCAGAAAGCTGGGCCCTTATCTGTGTTTTCGGCTAACAAAGGGTGGACGCTTC 1623
1691 TGGGGCCGCTCCACCTGACCGGAGAGAGGGCGGCTTTGGCCTCAGCGCTTCGGGGAGACT 1750
1624 CTCGAAGCATCCGCTTCACTGCAGAAAGAGGGGACTTGGGGTTTCACTTTGAGAGGGAACG 1683
1751 CGCCTGTCTCATCGCTCGTCAATTCCAGGAGAGCGCGCGGCTGGCGCTGAAGG 1810